

Readability and Suitability of Implantable Hearing Device Information Published Online in English

Matthew Goddard

Master of Audiology thesis (CMD690)

PRIMARY SUPERVISOR

Associate Professor Rebecca Kelly

SECONDARY SUPERVISOR

Associate Supervisor Megan McAuliffe

January 30, 2020



This thesis (CMDS690) is being submitted in partial fulfilment of University of Canterbury requirements for the degree, Master of Audiology (MAud).



Acknowledgements

First, I wish to thank my wife Liping and our two sons Han and Sebastian. Their patience and optimism always made me feel humble, especially during those busy days where time together as a family seemed to occur only during mealtimes.

Second, I would like readers to be aware of just how difficult a time my primary supervisor Rebecca has had over this past 24 months. I recall vividly Rebecca telling our cohort during a lecture that she was experiencing significant health issues; yet, during this very same semester, and without holding back, she offered to supervise all our thesis projects. During thesis-year Rebecca directly supervised seven of us (half the class). She kept us on track, reminded some of us on weekends to send her our Introduction to check, organised one-on-one and group meetings and statistics workshops—all while facing challenging internal department politics, heavy teaching duties, and tending to her own health. I mention all this because I wish the reader to know that Rebecca is easily the most kind and generous supervisor I have ever had.

Third, I recall back in 2017 something my mother said to me when I told her I was giving serious thought to applying for this course. She said: “That’s the smartest thing I’ve ever heard you say.” Turns out she had been waiting 38 years for that small epiphany. I’m glad it happened. And as with Liping and our boys, I’m forever grateful for the support of my whānau.

Abstract

The aim of this Master of Audiology thesis was to investigate online English-language information relating to implantable hearing devices, such as cochlear implants, and bone-anchored hearing aids. Specific goals included: (1) to assess the readability of materials (Flesch-Kincaid, SMOG, Gunning-Fog), and relate this to adult literacy levels; (2), to assess suitability of online information (DISCERN, PEMAT, Plain Language, SAM+CAM), and relate this to search engine queries people use in order to research such information; and (3), to test whether or not organisation type, geographic location, and HONcode certification act as moderating factors on webpage readability and suitability. Methods included: surveying key search terms related to “implantable hearing devices,” obtained using convenience sampling ($n=25$), followed by refinement of search terms using Google Trends. Results showed that webpages placed high literacy demands on readers, requiring on average a 10-12th reading grade level for basic comprehension, which far exceeds recommendations set by several organisations. Serious shortcomings were also found concerning incomplete discussion of treatment choices, transparency of information sources, confusing visual aids, and frequent use of technical terms without sufficient lay explanation. Multivariate and univariate statistical analyses revealed no significant effect of webpage location, organisation type, or HONcode certification on readability and suitability measures. Conclusions: several aspects of webpage content design act as barriers to effective information gathering. However, neither HONcode certification nor web domain location nor organisation type seem to be among them.

Contents

1	Introduction	1
1.1	Health literacy—what is it?	3
1.2	Implications of low health literacy	4
1.3	Health literacy in New Zealand	5
1.4	Assessing health related information	5
1.4.1	Readability	6
1.4.2	Suitability	8
1.4.3	Reliability	10
1.5	Implantable hearing devices	12
1.5.1	Cochlear implants	12
1.5.2	Bone conduction implants	14
1.5.3	Middle-ear implants	15
1.5.4	Auditory brainstem implants	15
1.5.5	What information exists for those interested in hearing implant devices?	16
1.5.6	Present thesis scope	16
1.6	Research aims	17
1.7	Hypotheses	18
1.8	Expected findings	18
2	Methods	20
2.1	Study participants	20

2.1.1	Recruitment	21
2.1.2	Instructions	21
2.1.3	Ethics approval	21
2.2	Units of analysis: webpages	21
2.2.1	Search terms	21
2.2.2	Web landing page	22
2.3	Webpage assessment	22
2.3.1	Readability	22
2.3.2	Suitability	23
2.3.3	Reliability	30
2.3.4	Inter-rater reliability	31
2.3.5	Statistical analyses	31
3	Results	33
3.1	Survey response	33
3.2	Search terms	35
3.3	Inter-rater agreement	36
3.4	Readability of webpages	37
3.4.1	Location	38
3.4.2	Organisation type	39
3.4.3	Interaction between organisation type and location	40
3.4.4	HONcode	41
3.5	Suitability of webpages	42
3.5.1	Location	43
3.5.2	Organisation type	44
3.5.3	Interaction between organisation type and location	46
3.6	Reliability: HONcode	47
3.7	Videos	49

4	Discussion	51
4.1	Webpage readability	51
4.2	Webpage suitability	54
4.3	Global differences in webpage readability and suitability . .	57
4.4	Commerical webpages	58
4.5	HONcode	59
4.6	Audiovisual materials	60
4.7	Limitations	61
4.8	Future recommendations	64
5	Conclusion	66

List of Figures

1.1	Current options for implantable devices by type, as categorised elsewhere (Tisch, 2017)	13
3.1	Counts of individual respondent age, bin size = 10 years . .	33
3.2	Mean RGL (\pm 95% CI) among webpages hosted by organisations located in the Americas (1), West Pacific (2), and World/Other (3). FK = Flesch-Kincaid, GF = Gunning Fog, SMOG = simple measure of gobbledegook. . .	39
3.3	Mean RGL scores (\pm 95% CI) among webpages hosted by for-profit (FP) and not-for-profit (NFP) organisations. FK = Flesch-Kincaid, GF = Gunning Fog, SMOG = simple measure of gobbledegook.	40
3.4	Mean RGL among webpages with nesting on organisation type (for-profit, FP; not-for-profit, NFP) and WHO region (Americas, 1; West Pacific, 2; World/Other, 3) Error bars = 95% CI Note: Mean RGL refers to mean across GF, FK, and SMOG; i.e., consistent with MANOVA analysis.	41
3.5	Suitability scores for PEMAT understandability (P-U), plain language (PL), and SAM+CAM (S+C) for webpages hosted by organisations located in the Americas (1), West Pacific (2), and World/Other (3). Error bars = 95% CI	44
3.6	Suitability scores for PEMAT understandability (P-U), plain language (PL), and SAM+CAM (S+C) for webpages hosted by for-profit (FP) and not-for-profit (NFP) organisations. Error bars = 95% CI	46

3.7	Interaction between WHO region and type of organisation for web page suitability. Mean RGL among webpages with nesting on organisation type (for-profit, FP; not-for-profit, NFP) and WHO region (Americas, 1; West Pacific, 2; World/Other, 3) Error bars = 95% C.I. Note: Mean suitability score (%) refers to mean across PEMAT (understandability), plain language, and SAM+CAM; i.e., consistent with MANOVA analysis.	47
-----	---	----

List of Tables

2.1	PEMAT for print/webpage material	24
2.2	PEMAT for audiovisual material	25
2.3	DISCERN tool scoring system	27
2.4	SAM+CAM inventory	28
2.5	Plain Language Checklist	30
3.1	Respondent gender, education, ethnicity	34
3.2	Respondent age, terms supplied, and time taken to complete survey	34
3.3	List of search terms refined from survey responses	35
3.4	Refined search terms for Google Trends analysis	36
3.5	Reading grade level for all webpages, excluding videos	37
3.6	Spearman Correlations	37
3.7	Marginal RGL means - WHO region	38
3.8	Marginal RGL Means - Organisation Type	40
3.9	Marginal RGL means - Location x Organisation Type	41
3.10	Suitability scores for all webpages	42
3.11	Spearman correlations between suitability inventories	43
3.12	Marginal means (%) - Suitability by WHO region (SAM+CAM and PEMAT (understandability)	44
3.13	Marginal means (%) - Suitability by organisation type (excluding PEMAT (actionability) and DISCERN)	45

List of abbreviations

ANOVA, analysis of variance

BAHA, bone anchored hearing aid

FK, Flesch-Kincaid

FP, for-profit organisations

GF, Gunning's Fog Index

GP, general practitioner

HONcode, health on the internet code of conduct

ICC, intraclass correlation coefficient

IQR, interquartile range

MANOVA, multivariate analysis of variance

MoH, Ministry of Health

NFP, not-for-profit organisations

NZ, New Zealand

NAAL, national assessment of adult literacy

NALS, national adult literacy survey

NIH, national institutes of health

M, arithmetic mean

Mdn, median

PEMAT, patient education materials assessment tool

PL, plain language

PLAIN, plain language action and information network

SMOG, simple measure of gobbledegook

CI, 95% confidence interval

REALM, rapid estimate of adult literacy in medicine

RGL, reading grade level

SE, standard error the mean

TOFHLA, test of functional health literacy in adults

U.K., United Kingdom

U.S., United States of America

WHO, world health organisation

WRT, wide range achievement test

Chapter 1

Introduction

In not quite three decades the World Wide Web has grown from approximately 10 webpages to over 1.6 billion (InternetLiveStats, 2018; Berners-Lee et al., 2010). Today more than half the global population accesses the Internet (IWS, 2019). Among the multitude of reasons for doing so, gathering health information appears to be of growing importance. Prior surveys carried out between 2000 to 2012 suggest in the order of 50–80% of U.S. adults search the Internet for health information; either because of healthcare access barriers; or for self-efficacy reasons e.g. healthier lifestyle, improving patient-physician interaction; or because the intended benefit is for someone else, e.g. family member (Fox S, 2012; Ingram, 2015; Kim et al., 2017; Amante et al., 2015; Chu et al., 2017; Sillence et al., 2007). A survey completed in November 2005 involving seven European countries (Denmark, Germany, Greece, Latvia, Norway, Poland, Portugal) found 44% of the 7,934 participants interviewed used the internet for health purposes; 29% of whom so did in order to decide whether or not to see a GP (Andreassen et al., 2007).

The means by which people obtain health information has also seen change. Smartphones and tablets surpassed traditional computers (desktops, laptops) in 2013 for online health queries, at least concerning web traffic to Mayo Clinic; and in particular for queries related to “symptoms” (Jadhav et al., 2014). Further to this, with smart device penetration estimated at over 90% 2013 global population, it is without

surprise that today in 2019 this type of technology is relied upon heavily for seeking online health information IWS (2019).

According to Laurent and Vickers (2009) health information is increasingly accessed through decentralised or “open-source” mediums. Analysis of Google webpage rankings for common health-related queries revealed Wikipedia ranked within top 10 results in 75% of searches; rivalling government domains (i.e. “.gov” uniform resource locator (URL) suffix), and surpassing popular web health portals such as MayoClinic.org and WebMD.com and KidsHealth.org, among many others (Laurent and Vickers, 2009). In keeping with this, studies have found that establishing trust ranks high amongst consumers needs. To this end, information needs to be relevant, unbiased (i.e., sole objective is not to sell a product or service), and from a knowledgeable source with reasoning clearly explained (Sillence et al., 2007; Toms and C, 2007; Glenton et al., 2005).

In some respects these basic requirements are unsurprising in light of why people choose to seek out health information in the first place. Autonomy in decision-making is one key driver (e.g., as to whether or not to seek out medical help in the first place); so is the need to gain mastery over a given health topic, i.e. before and/or after GP consultation (Santana et al., 2011; Andreassen et al., 2007). Other key motivators include reassurance-seeking over proposed treatments and diagnoses, and information-seeking on possible lifestyle and dietary changes, presumably also following a diagnosis or treatment (Wangberg, 2009).

Of course, none of this should matter if the very information people access cannot be easily understood. Herein lays an important caveat: more information can, in some cases, or perhaps many cases, amount to less. Educating people so that they may better understand technical concepts is one way of overcoming this problem, but it is unreasonable to expect every person avail themselves of a medical or medical-science degree just so that they may weigh up risks of various antibiotic treatments on hearing loss, for example. A more natural solution is to target the content itself and render it as easily understood as possible. A question one might ask is just how understandable does content need to be. This question is addressable through *health literacy* research.

1.1 Health literacy—what is it?

Health literacy is essentially knowledge of health information. Formal definitions have evolved over time, including how it should be measured (Berkman et al., 2010). During early 1990s, United States Department of Education held health literacy to mean: “an individual’s ability to read, write and speak in English, and compute and solve problems at a level of proficiency necessary to function on the job and in society, to achieve one’s goals, and develop one’s knowledge and potential” (as quoted in (Kirsch et al., 1993, p.3)). From this— and in an effort to capture the state of U.S. health—a National Adult Literacy Survey (NALS) undertook to show health burden related to poor literacy. Subsequently in 2003, National Assessment of Adult Literacy (NAAL) in reporting specifically on “health literacy” and using the same operational definition of literacy as Kirsch et al. (1993), showed a positive relationship between health literacy score and overall health (Kutner et al., 2003). Specifically, that: education positively influenced health literacy proficiency; and that low health literacy and poverty were related, as was literacy and the degree to which health information was accessed in books and online; and that health literacy varies by occupation, for instance those in farming/fishing/forestry tending towards lowest health literacy compared higher scoring professional vocations.

Conceptual refinement during early 2000 subsequently held health literacy to mean “the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” (as quoted in (Selden et al., 2000) p.vi). At the most basic level, therefore, individuals must be able to at least read health information if they are to participate in health decision-making, e.g., to follow printed instructions on dosing regime. Nutbeam (2000) referred to this as *basic* or *functional* literacy. Greater autonomy in health decision-making, e.g., to decide among treatments, requires higher-level cognitive and social skills—higher still if one expects to suggest an appropriate treatment course of action to their healthcare practitioner. Accordingly, these fell under two additional classifications:

interactive or *communicative* literacy, and *critical* literacy, respectively (Nutbeam, 2000).

Under such frameworks health literacy can be seen as a risk factor for adverse health outcomes through which targeted health education serves to mitigate. It follows that any improvement to how easily health information may be presented and detailed should translate to greater self-efficacy and ultimately improved clinical outcomes. Alternatively, health literacy can be thought of as an *outcome* of education and communication, which, as Nutbeam (2008) suggested, may be favourable in public health because health literacy is then seen as an asset to be invested in at the individual level rather than a risk to be modified at organisational or societal levels.

In line with the latter model, Berkman et al. (2010) suggested health literacy be thought of as “The degree to which individuals can obtain, process, understand, and communicate about health-related information needed to make informed health decisions”—thus, emphasizing action and communication, while respecting that some individuals may well like to understand more than just “basic health information.” This latter definition is consistent with New Zealand Ministry of Health/Manatū Hauora (MoH), namely: “Health literacy is the capacity to find, interpret and use information and health services to make effective decisions for health and wellbeing” (MoH, 2019).

1.2 Implications of low health literacy

Adults with low functional health literacy have been known to hide their true level of understanding by pretending to understand scripts and other health information out of shame (Parikh et al., 1996). At the very least this creates barriers to optimal care. At worst, individuals place themselves at greater risk of health decline. As one recent systematic review found, poor reading ability¹ increases the risk 1.2 to 1.5-fold that a

¹Wide Range Achievement Test (WRAT); Rapid Estimate of Adult Literacy in Medicine (REALM); Test of Functional Health Literacy in Adults (TOFHLA)

person will experience adverse health outcomes, including hospitalisation (Dewalt et al., 2004). Contrariwise, those with adequate health literacy are more likely to comply with medicinal prescriptions in support of better health outcomes (Lee et al., 2017).

1.3 Health literacy in New Zealand

According to MoH, low health literacy results in underutilisation of preventative and screening services. As a consequence, New Zealanders with low health literacy may inadvertently exacerbate their chronic conditions to the point of requiring hospitalisation. In addition, because people with low health literacy lack understanding of their health conditions, they also lack knowledge of treatment choices and medicines, and fail to properly manage their chronic conditions (MoH, 2010).

Health-literacy inequalities between Māori and non-Māori also exist. MoH statistics show 75–80% of Māori possess low health literacy skills, compared to 60% for non-Māori. New Zealanders’ level of education also has an effect on health literacy, with those achieving only minimal secondary school education being unable to adequately understand health information presented to them. Furthermore, of Māori who complete higher secondary education, 90% show poor health literacy skills, compared to 60% of non-Māori. While at the tertiary level, only 50% of Māori show good health literacy skills, compared to 66% of non-Māori (MoH, 2010).

1.4 Assessing health related information

This thesis project is concerned with applying four broad constructs to assess how well individuals might access and potentially comprehend hearing-implant-device related information. In the Nutbeam (2008) model of health literacy this corresponds to the level of “tailored health information, communication, education” (i.e., Figure 1 Nutbeam (2008)

p.2074). It is the second-lowest level in the health literacy model above that of the individual level; i.e., reading ability, numeracy and general knowledge. This distinction is made only to clarify that the goal herein is not to measure health literacy *per se*. Rather, the goal is measure one of several fundamental components for which health literacy is derived.

1.4.1 Readability

Understanding literacy demands placed on readers requires complexity-based statistics of text (Zamanian and Heydari, 2012). Early implementation of this idea used word frequency analysis after taking into account reader age, i.e., The Teacher’s Word Book by Thorndike (1921)¹. One popular² mid-twentieth century implementation of readability by Flesch (1948) used a more sophisticated approach by taking into account four elements: average sentence length in words, average word length in syllables, percentage of personal pronouns, and percentage of sentences using quotation marks and punctuation used to engage or address the reader directly. From this, a measure of reading ease was calculated, and using regression analysis grade level could be estimated for the text. This approach was simplified by Kincaid et al. (1975) to Eq. (1.1):

$$GradeLevel = 0.39 \cdot wps + 11.80 \cdot spw - 15.59 \quad (1.1)$$

where, wps = average number of words per sentence; spw = average number of syllables per word. The Flesch-Kincaid (FK) method of estimating grade level is one method utilised in the present thesis to assess text readability.

Other established readability formulae used in this project include the Simple Measure of Gobbledegook (SMOG) and Gunning’s Fog Index (GF). The SMOG readability formula was introduced by McLaughlin in 1969 as a tool which the author passionately advocated as:

¹Edward Thorndike extensively catalogued an alphabetical list of ten-thousand English words, out of 625,000 words common in children’s literature, and published these in a 150-page manual for teachers. An electronic out-of-copyright reproduction of the 1921 manual can be found at: <https://archive.org/details/teacherswordbook00thoruoft>

²3,967 citations according to Google Scholar as at April 2019

“ ... laughably simple, it is in fact more valid than previous readability formulas,” (Mc Laughlin, 1969, p.639).

McLaughlin cites the work of Gunning whose own readability formula takes into account the number of polysyllabic words within a 100-word body of text in order to predict reading difficulty (Gunning, 1952). SMOG uses the same principle. The difference being, McLaughlin recommended counting the number of polysyllabic words over 30 sentences (approximately 600 words) because this offered greater predictive power of reading grade level (RGL). Using regression analysis (student grade score vs. standardised text polysyllable count), McLaughlin then simplified SMOG to Eq. (1.2):

$$GradeLevel = 3 + \sqrt{pwps} \quad (1.2)$$

where, pwps = the number of polysyllable words in 30 sentences (Mc Laughlin, 1969);

Gunning’s Fog index was developed by Robert Gunning in the late 1940s through 1952 (Gunning, 1952). It was intended as a guide for assessing and subsequently improving clarity and ease of reading of printed materials (e.g., newspapers, business articles). Gunning’s reference to “fog” (fogginess) was a metaphor for lack of clarity in commercial and government writings. It was also a euphemism, as later revealed within the Journal of Business Communications: “Through pomposity, stupidity or carelessness they use more elaborate signals, from sentence to sentence, than their messages require” (Gunning, 1969, p.11). The formula takes into account sentence length and the proportion of difficult words (see Eq. (1.3) below) before arriving at an index. This is the estimated number of school-years required to understand English language text.

$$GradeLevel = .4 \cdot (awps + psw) \quad (1.3)$$

where, awps = average words per sentence of a passage of 100 or more words; and psw = percentage of words three syllables or more, excluding inflected verbs (ending in “ing” or “ed” or “es”) (Gunning, 1952).

1.4.2 Suitability

Goals of suitability assessment differs markedly from that of readability. For instance, unlike readability formulas, suitability assessment tools take into account content flow and style, how personable it is made, whether and to what degree landing pages bear relevance to the preceding search query, and so on. In this way, suitability tools are particularly well suited for determining overall webpage relevance and quality; which, for consistency with other studies, is referred to here as “suitability.”

Below, four popular suitability tools are introduced. All of which were used in this thesis. These are: DISCERN, SAM+CAM, PEMAT, and PL.

DISCERN is a free to use tool developed in 1998 for content producers in the hope that, through careful content design, patients and consumers are able to make informed choices from quality evidence-based information (Charnock, 1998; Charnock et al., 1999). DISCERN consists of a 16 item questionnaire, scored using a five-point likert scale. Scoring an item “1” indicates no fulfilment of the quality criteria; scores of “2” “3” ‘ or ‘4” indicate partial fulfilment; and “5,” complete fulfilment. The questionnaire is divided into three themes (see Section 2.3.2.2). The first of which identifies whether or not sources of information can be deemed credible, and how well the information relates to the aim of the webpage. Section two specifies seven questions relating to treatment choices. A third section, with a single question, asks the rater to assign an overall quality score for the webpage as a source of information about treatment choices.

SAM+CAM is a 22-item instrument developed in 2009 as a modification of the Suitability Assessment of Materials (SAM) instrument, introduced over a decade earlier by Doak et al. (1996). The original SAM tool was developed to fulfil an unmet need among healthcare organisations, who otherwise wished to quickly assess the suitability of patient education materials, but lacked sufficient resources to carry out this process. As with SAM+CAM, SAM contains 22 items that assess content scope and purpose, literacy demand including RGL, use of illustrations, layout and typography, motivation to attend key messages, and cultural awareness. As can be seen from Table 2.4 many of these items remain in

SAM+CAM. Where the two tools differ is with the inclusion of numeric literacy assessment in SAM+CAM, as well as assessments of persuasive techniques (e.g., appealing to familiar values), application of learning theory (e.g., “active learning,” see Michael (2006)), and communication assessment (e.g., addressing reader directly, use of analogies in explanations) (Helitzer et al., 2009). Of particular relevance to this thesis is numeric assessment. This is because mathematical and engineering Standard International units frequently appear in the context of implantable hearing devices. One additional point of difference worth mention is that since SAM+CAM does away with RGL, it is not only faster but is less redundant—since most health literacy studies tend to use a combination of readability formulae with suitability tools.

PEMAT is a content suitability tool that was introduced in 2014 specifically for assessments surrounding how well content may be understood, and how well content prompts readers to take action. These are referred to as “understandability” and “actionability,” and consist of 19 and 7 questions, respectively; each requiring an agree/disagree response. Shoemaker et al. (2014, p.396) defines each of these as follows:

“Understandability: Patient education materials are understandable when consumers of diverse backgrounds and varying levels of health literacy can process and explain key messages.”

“Actionability: Patient education materials are actionable when consumers of diverse backgrounds and varying levels of health literacy can identify what they can do based on the information presented.”

One clear advantage PEMAT has over other suitability tools is that it can be applied to print as well as audiovisual content, such as YouTube (Shoemaker et al., 2013c,b,a, also see Tables 2.1 and 2.2). Another useful characteristic is the degree of overlap shared with SAM+CAM in terms of numeracy assessment, and webpage aims as with DISCERN. Thus, when used together, additional validity can be gained through cross-checks of similar-item ratings.

The fourth suitability tool, Plain Language, is actually the newest of all four herein discussed because it was developed in-house by two Master of Audiology thesis students. It is based on two existing checklists: National Adult Literacy Agency (NALA) Checklist for Documents, and Plain Language Action and Information Network (PLAIN) Checklist for Plain Language on the Web (Greene et al., 2017).

1.4.3 Reliability

Estimating reading grade level together with at least one content suitability analysis tool is the approach preferred by many recent studies that assess medical or health related patient education materials (e.g., Alamoudi and Hong, 2015a; Boston et al., 2004; Greer, 2019; Grewal and Alagaratnam, 2013a; Harris et al., 2018; Joseph et al., 2016; Ritchie et al., 2016; Seymour et al., 2015a). One additional quality indicator not mentioned thus far takes an entirely different approach—HONcode certification. This is an initiative by Health On The Net Foundation, a non-governmental organisation that aims to safeguard quality and transparency of health information for patients, clinicians, and the general public. To date over 8,000 websites subscribe to HONcode certification (HONcode, 2020). Medical website publishers can choose to apply for HONcode certification to show that their information is objective, of sufficient quality, and transparent in intention. In essence, reliable (trustworthy). In so doing, web domain administrators must adhere at all times to a set of principles (see below, quoted from: HONcode, 2020). Although, as studies have found, not all website administrators adhere to all eight HONcode principles (Boyer et al., 2011; de Castro Correa et al., 2013).

1. “Authoritative: Any medical or health advice provided and hosted on this site will only be given by medically trained and qualified professionals unless a clear statement is made that a piece of advice offered is from a non-medically qualified individual or organisation.”

2. “Complementarity: The information provided on this site is designed to support, not replace, the relationship that exists between a patient/site visitor and his/her existing physician.”
3. “Privacy: Confidentiality of data relating to individual patients and visitors to a medical/health Web site, including their identity, is respected by this Web site. The Web site owners undertake to honour or exceed the legal requirements of medical/health information privacy that apply in the country and state where the Web site and mirror sites are located.”
4. “Attribution: Where appropriate, information contained on this site will be supported by clear references to source data and, where possible, have specific HTML links to that data. The date when a clinical page was last modified will be clearly displayed (e.g. at the bottom of the page).”
5. “Justifiability: Any claims relating to the benefits/performance of a specific treatment, commercial product or service will be supported by appropriate, balanced evidence in the manner outlined above in Principle 4.”
6. “Transparency: The designers of this Web site will seek to provide information in the clearest possible manner and provide contact addresses for visitors that seek further information or support. The Webmaster will display his/her E-mail address clearly throughout the Web site.”
7. “Financial disclosure: Support for this Web site will be clearly identified, including the identities of commercial and non-commercial organisations that have contributed funding, services or material for the site.”
8. “Advertising policy: If advertising is a source of funding it will be clearly stated. A brief description of the advertising policy adopted by the Web site owners will be

displayed on the site. Advertising and other promotional material will be presented to viewers in a manner and context that facilitates differentiation between it and the original material created by the institution operating the site.”

HONcode certification status can be checked either through the Health On The Net Foundation, or one of three freely available web browser extensions for Firefox, Chrome, and Internet Explorer. Efforts are being made to accelerate certification detection using automated search engines, although at present manual detection of all eight principles remains more accurate (Boyer and Dolamic, 2015).

1.5 Implantable hearing devices

This thesis has specific interests in the readability, quality, and reliability (trustworthiness) of English-language webpages concerned primarily with implantable hearing device technology. Even though this field remains somewhat niche, it consists of several well-established technologies: active middle-ear implants, bone conduction implants, and cochlear implants (for a review, see Tisch, 2017). A fourth and somewhat more radical technology is that of auditory brainstem implants (see, Nakatomi et al., 2016). Fig. 1.1 provides a summary of these technologies.

1.5.1 Cochlear implants

Cochlear implants predate all aforementioned implantable hearing device technology by more than two decades (Macherey and Carlyon, 2014; Tisch, 2017). The role of CI devices is to convert sound pressure waves into digital signals that can then be fed directly to the cochlea—the organ of hearing. In this way the CI stands in for most of the work performed by the peripheral auditory system: i.e., microphones housed in the external part of the CI collect sound pressure waves transmitted through air, ordinarily the job of the pinna, external auditory canal and tympanic

IMPLANTABLE HEARING DEVICES

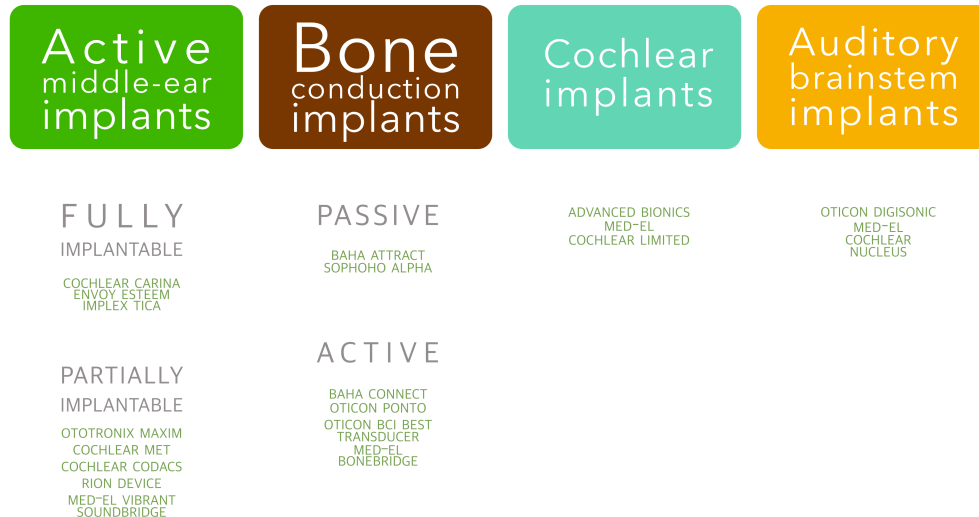


Figure 1.1: Current options for implantable devices by type, as categorised elsewhere (Tisch, 2017)

membrane. A digital signal processor also housed in the external CI along with a battery then converts the signal into patterns of electrical pulses, ordinarily the job of the cochlea inner hair cells; the resulting patterns of electrical energy are then delivered via the internal part of the CI—the electrode array—to nerve fibres that form the auditory nerve (cranial nerve VIII). Because this relies on adequate functional cochlea regions which at time of surgery cannot precisely be established³, hearing outcomes among CI patients can be highly variable (Macherey and Carlyon, 2014), with younger patients generally able to perform better with speech testing than older patients (>60 years)—although there is some indication that hearing-related quality of life is similar across ages (Sladen and Zappler, 2015).

Early U.S. Food and Drug Administration estimates held that 324,000 cochlear implant devices had been implanted worldwide as of 2012 (NIH,

³Technically this would be possible with psychometric tuning curves in adults and young people able to behaviourally respond to sound stimuli (Moore and Glasberg, 1997; Moore and Alcántara, 2001). For infants and toddlers, however, finding cochlea dead regions using psychometric tuning curves is not practical.

2019). New Zealand presently funds approximately 100 implants per year, up from 20 per year in 2010 (Gunn, 2010), which is split between the Northern Cochlear Implant Program (NCIP) and Southern Cochlear Implant Program (SCIP).

1.5.2 Bone conduction implants

Unlike CIs and ABIs which interface directly with neural tissue, the remaining two classes of implantable hearing devices (BC and ME implants) both rely on intact and functioning cochleae. BC implants take advantage of the fact sound energy is transmissible through fluid within the cochlea by surrounding cranial bone—a phenomenon first reported in sixteenth century otology by Capivacci (Kelley, 1937).

By way of clarification, two broad types of bone-conducting hearing aids are presently available—those anchored surgically to cranial bone (i.e., bone-anchored hearing aids, or BAHA) and those that do not require surgery. The latter, more commonly referred to as bone-conducting hearing aids (or BCHA) predate BAHAs and work by the same mechanism, although less efficiently (Dobrev et al., 2017). Because they are not implantable devices, however, BCHAs will take no further mention here.

Normal indications for BAHAs include those with inoperable congenital defects of the outer ear in which the ear canal is absent or closed (aural atresia), hearing loss with (or as a result of) chronic discharge (chronic otitis media), and those with acquired or congenital single-sided deafness (Snik et al., 2004). The goal in each case is to re-route sounds from the side of hearing impairment to that of the good ear, thus restoring speech cues and sounds ordinarily attenuated or lost through “head shadowing.” BAHAs are generally well tolerated with few complications, and provide favourable audibility outcomes compared to surgical reconstruction, e.g., canaloplasty, tympanoplasty, and stapes and ossiculoplasty (Bento et al., 2012; Saroul et al., 2011).

1.5.3 Middle-ear implants

As shown in Figure 1.1 (above), middle ear implants fall under two categories distinguished on the basis of where the digital sound processor and battery sit. Fully implantable devices like the Esteem (Envoy Medical) and Carina (Cochlear) represent the most discrete hearing device option currently on the market, because 100% of the device’s components are concealed beneath the skin—including the microphone. Partially implantable devices like the Maxum (Ototronix) and MET (Cochlear) house their digital signal processor externally, e.g., in the ear canal or behind the ear. Advantages and disadvantages in terms of surgical complexity and cost are thoroughly explained in a recent review by Tisch (2017). Regardless of whether the middle ear device is fully or partially implantable, all share in common the basic mechanism of sound delivery to the cochlea via the ossicular chain. For this reason, indications include those with sensorineural hearing loss in which other treatment interventions have proved ineffective. However, as reviewed by Lassaletta et al. (2019) the past decade has seen important advances to middle ear implant technology (e.g., direct mechanical stimulation of the oval window) such that present indications now include mixed conductive and sensorineural hearing loss, with or without anatomically normal middle ear structures.

1.5.4 Auditory brainstem implants

As mentioned, this is a somewhat more radical and in some ways exciting “brain machine interface” type of technology that relies on no peripheral auditory structures. That said, auditory brainstem implants in effect share much in common with cochlear implants. Both consist of external speech processors that transduce sound pressure waves to trains of electrical current pulses, which are then delivered to nerve tissue. Crucially however, where auditory brainstem implants differ is in the site of stimulation: cochlear implants stimulate auditory nerve primary afferents (auditory division of cranial nerve VIII); auditory brainstem

implants stimulate brainstem cochlea nuclei. Indications for the latter are specific to cases where auditory nerve pathways to the brainstem are bilaterally impaired; e.g., from cerebellopontine angle tumours or other cranial pathology (see Table 1., Nakatomi et al., 2016). To date, seven-hundred or so procedures have reportedly been carried out since 1979. Clinical outcomes range from 4% to 77% improvement in word recognition performance, and 0% to 31% improvement in sentence recognition performance (Nakatomi et al., 2016).

1.5.5 What information exists for those interested in hearing implant devices?

Given what complexities lay ahead for information seekers, particularly in cases where ME implants are being considered as a treatment option, access to suitable information early on has tremendous and lasting impact on quality of life. For this reason it is imperative that patients are aware of their treatment options, including the risks, costs, and pros and cons each technology has to offer. Unfortunately, as Seymour et al. (2015a) found, information written about cochlear implant devices generally exceeds reading levels for most adults. This trend is not unique to implantable devices. As Laplante-Lévesque and Thorén (2015) found in their systematic review, FK reading grade levels of 10–14 are not uncommon within the scope of myringotomy, acoustic neuroma, and hearing impairment in adults. Admittedly the Seymour et al. (2015a) study is several years old now and in that time technology and health information likely has evolved. Notwithstanding, very little is known about readability and suitability of implantable hearing device technology and their respective surgical procedures, and it is on this basis that the current thesis project draws inspiration.

1.5.6 Present thesis scope

Access to simple-to-understand information for patients and families is crucial in empowering people to make informed decisions when faced with

various implantable hearing technology treatment options. Research has already shown that patients immediately forget at least one-quarter the information presented to them during pre-surgical consultation (Feiner and Rayan, 2016). This implies that independent research, in principle, constitutes an equally vital source of information alongside that provided by clinicians. If, however, the very information being accessed is written at too advanced a level, or is factually incorrect, outdated, incomplete, misleading, biased, or irrelevant—then there is every chance for decision-making to be rendered suboptimal. Added to that:

- over half the world’s population has access to the internet (IWS, 2019);
- online content grows daily, not necessarily in quality but in quantity;
- World Health Organisation statistics hold that some 466 million people worldwide currently experience a form of disabling hearing loss which is anticipated to grow to 900 million by 2050 (WHO, 2019).

In light of the above, this Master of Audiology thesis proposal outlines a project in which the current state of implantable hearing device web-information is given closer examination.

1.6 Research aims

This project’s overall aim is to evaluate how useful existing sources of internet information is for people wishing to learn more about implantable hearing devices. Following are specific research questions:

1. Do the literacy demands of most “implantable hearing device” themed webpages enable people to make informed health decisions, or do they present barriers to acquisition of information?
2. How relevant and helpful are these webpages? Is the information of high quality and free from bias?

3. Are webpages hosted by for-profit organisations any different from not-for-profit?
4. What differences, if any, might emerge if webpages from different English-speaking countries are compared?

1.7 Hypotheses

From the above research questions, the following formal hypotheses were tested. The null hypothesis was defined as:

H_0 : readability and suitability measures will not differ significantly by: (a) WHO region of hosting organisation, (b) organisation type, (c) WHO region by organisation type, (d) HONcode certification.

The alternative hypothesis tested against the null was therefore:

H_1 : information on implantable hearing devices will differ significantly in terms of readability and suitability, depending on: (a) location of hosting organisation, (b) type of hosting organisation, (c) HONcode certification, (d) location and organisation type.

1.8 Expected findings

In undertaking this thesis project the following outcomes were anticipated:

1. Overview of top-ranking webpages related to the present theme, in terms of readability and suitability
2. Evidence demonstrating that regional differences affect readability and suitability
3. Similarly, evidence demonstrating whether or not organisation type affects readability and suitability

4. Finally, evidence demonstrating whether or not HONcode certification influences readability and suitability

Chapter 2

Methods

2.1 Study participants

To generate an initial list of key search engine terms, study participants were recruited using convenience. Among those targeted for recruitment were friends of the student researcher, family members, school teachers, friends of friends, and personal referrals within the Christchurch Public Hospital. Compensation (financial, gift vouchers, or otherwise) was not offered to participants involved in this study. Inclusion criteria were as follows:

- Fluency in the English language
- Male, female, gender diverse, adults over the age of 18 years who indicated their consent to participate
- Any ethnicity
- Must be able to supply at least three search terms
- Internet access

Exclusion criteria were: not fluent in the English language (survey would not progress if this was true), consent not indicated, age \leq 18 years.

2.1.1 Recruitment

Email served as the primary medium for recruiting study participants. Private Facebook messenger chats were also used. Potential study participants who clicked the provided URL were taken to a web landing page where they could read more about the study by way of downloadable Survey Brief (see Section 3.6 for a copy). Informed consent was obtained by clicking on an appropriately labelled button on the landing page; clicking this allowed participants to begin the survey. Recruitment continued until saturation, defined as: >2 search term repeats or near-repeats in succession or close succession (i.e., within 2–3 consecutive surveys completed).

2.1.2 Instructions

Landing page instructions consisted of asking potential study participants to read through the Survey Brief, and, should they so wish, provide consent. Each consenting participant therefore understood the following information would be collected: gender, highest level of education, ethnicity, age, search terms (≥ 3 required).

2.1.3 Ethics approval

Approval for this project has been granted under HEC 2019/07/LR (see Section 3.6)

2.2 Units of analysis: webpages

2.2.1 Search terms

Anonymous raw survey data was exported from Qualtrics Survey using in-built administrator tools. From here data were imported into Numbers (v.6 MacOS, Apple). Preliminary search terms filtered out if deemed

unsuitable or unrelated to the topic (Section 3.2). Search terms were then grouped and ranked by frequency. Groupings were assigned manually on the basis of common root words (inflection, spelling errors both ignored), synonyms (e.g., “cost” and “expense”), and common themes (e.g., surgical procedures, financial considerations, device information, physiology, deaf awareness).

Statistics were compiled Using Google Trends (worldwide, past 5 years). Low volume search terms were omitted. U.S./U.K. English spelling variants were queried separately. At this stage search terms were final.

2.2.2 Web landing page

From a list of final search terms, consecutive Google search queries were run using Mozilla Firefox (66.0.5) with a clean cache and browsing history and with all prior cookies removed. Top 10 search query hits (i.e., landing pages) provided the basic unit of analysis used in this study; these were referred to as “candidate pages,” which were further refined based on the following inclusion and exclusion criteria: (1) terms contained information relevant to implantable hearing devices; (2) landing pages included at least 100 words (or if a video, transcripts of at least 100 words); (3) landing page was freely accessible (not behind a paywall); (4) landing page was not a duplicate of a previously-included webpage (e.g., using a different search term).

2.3 Webpage assessment

2.3.1 Readability

Assessment of literacy demand was made using three widely cited readability formulas. As introduced in Section 1.4.1, these included Gunning Fog (GF), Flesch Kincaid (FK), and Simple Measure of Gobbledegook (SMOG). All three readability formulae were calculated using automated online software (Online-Utility.org, 2009). This required

a minimum 100-word body of text, consistent with inclusion criteria defined in Section 2.2.2. RGL calculations were obtained separately for each webpage. GF, FK, and SMOG data were copied and pasted into Numbers (MacOS, 10.15.2) for data storage and management.

2.3.2 Suitability

2.3.2.1 PEMAT

PEMAT scoring was based on 24 agree/disagree response type questions for print-based (Shoemaker et al., 2013b) and 17 agree/disagree response type questions for audiovisual-based (Shoemaker et al., 2013a) patient education materials. Each item was worth a total of 1-point (“agree”=1, “disagree”=0), for a total of 24 or 17 points depending on which version of PEMAT was used. Non-applicable questions were omitted from percentage calculations. By convention, rater responses marked as “agree” (and therefore scored 1) indicated that a characteristic was widely present in the material (i.e., 80-100%). Characteristics needing improvement or absent from material were marked “disagree” (and therefore scored 0). Tables 2.1–2.2 list characteristics for print-based and audiovisual-based PEMAT assessments (respectively: PEMAT-P, PEMAT-AV).

Table 2.1 PEMAT for print/webpage material

Item	Characteristic	Score ¹
CONTENT ²		
1	Material makes its purpose completely evident	_____
WORD CHOICE AND STYLE ²		
2	Material does not include information or content that distracts from its purpose	_____
3	Material uses common, everyday language	_____
4	Medical terms are used only to familiarise audience with the terms. When used, medical terms are defined	_____
5	Material uses an active voice	_____
USE OF NUMBERS ²		
6	Numbers appearing in the material are clear and easy to understand	_____
7	Material does not expect the user to perform calculations	_____
ORGANISATION ²		
8	Material breaks or “chunk” information into short sections	_____
9	Material’s sections have informative headers	_____
10	Material presents information in a logical sequence	_____
11	Material provides a summary	_____
LAYOUT AND DESIGN ²		
12	The material uses visual cues (e.g., arrows, boxes, bullets, bold, larger font, highlighting) to draw attention to key points	_____
USE OF VISUAL AIDS ²		
13	Material uses visual aids whenever they could make content more easily understood (e.g., illustration of healthy portion size)	_____
14	Material’s visual aids reinforce rather than distract from the content	_____
15	Material’s visual aids have clear titles or captions	_____
16	Material uses illustrations and photographs that are clear and uncluttered	_____
17	Material uses simple tables with short and clear row and column headings	_____
ACTIONABILITY		
18	Material clearly identifies at least one action the user can take	_____
19	Material addresses the user directly when describing actions	_____
20	Material breaks down any action into manageable, explicit steps	_____
21	Material provides a tangible tool (e.g., menu planners, checklists) whenever it could help the user take action	_____
22	Material explains how to use the charts, graphs, tables, or diagrams to take actions	_____
23	Material uses visual aids whenever they could make it easier to act on the instructions	_____

¹ Scoring: 0=disagreement, 1=agreement.

² Items one through 13 encompass understandability (PEMAT-U).

³ Actionability items span 14 through 17 (PEMAT-A).

Table 2.2 PEMAT for audiovisual material

Item	Characteristic	Score ¹
CONTENT ²		
1	The material makes its purpose completely evident	_____
WORD CHOICE AND STYLE ²		
2	The material uses common, everyday language	_____
3	Medical terms are used only to familiarise audience with the terms. When used, medical terms are defined	_____
4	Material uses an active voice	_____
ORGANISATION ²		
5	The material breaks or "chunks" information into short sections	_____
6	The material's sections have informative headers	_____
7	The material presents information in a logical sequence	_____
8	The material provides a summary	_____
LAYOUT AND DESIGN ²		
9	The material uses visual cues (e.g., arrows, boxes, bullets, bold, larger font, highlighting) to draw attention to key points	_____
10	Text on screen is easy to read	_____
11	The material allows the user to hear the words clearly (e.g., not too fast, not garbled)	_____
USE OF VISUAL AIDS ²		
12	The material uses illustrations and photographs that are clear and uncluttered	_____
13	The material uses simple tables with short and clear row and column headings	_____
ACTIONABILITY ³		
14	The material clearly identifies at least one action the user can take	_____
15	The material addresses the user directly when describing actions	_____
16	The material breaks down any action into manageable, explicit steps	_____
17	The material explains how to use the charts, graphs, tables, or diagrams to take actions	_____

¹ Scoring is out of two, with a score of zero indicating disagreement, and a score of one for agreement.

² Items one through 13 encompass understandability (PEMAT-U).

³ Actionability items span 14 though 17 (PEMAT-A).

2.3.2.2 DISCERN

Suitability ratings using the DISCERN tool were applied from Charnock (1998). DISCERN contains 16 questions divided in three sections. The first of which covers content reliability, and information about treatment choices. A final category, overall rating, contains a single item in which the rater assigns an overall quality score for the publication. Each question is scored on five-point scale, with scores of 1/5 through 5/5, respectively, for complete disagreement through complete agreement, with degrees of partial agreement in-between. Table 2.3 lists all 16 questions.

Table 2.3 DISCERN tool scoring system

Question	Characteristic	Score ¹
SECTION 1: CONTENT RELIABILITY		
1	Are aims clear?	_____
2	Does webpage achieve its aims?	_____
3	Is it relevant to users' needs?	_____
4	Are sources/references given?	_____
5	Is it clear how old sources of information is to user?	_____
6	Does the information seem unbiased and balanced?	_____
7	What additional sources of information are provided?	_____
8	Are areas of uncertainty admitted?	_____
SECTION 2: TREATMENT CHOICES		
9	Does it describe how each treatment works?	_____
10	Are benefits of each treatment outlined?	_____
11	Are risks of each treatment outlined?	_____
12	Does it explain what would happen in absence of treatment?	_____
13	Is quality of life mentioned for each treatment choice?	_____
14	How clear is it that ≥ 1 treatment choice exists?	_____
15	Does it provide support for shared decision-making?	_____
SECTION 3: OVERALL RATING		
16	Based on the answers to all the above questions, rate the overall ² quality of the publication source of information about treatment choices	_____

¹ Scoring based on five points in total. Where rater strongly agreed with characteristic, item was scored five. A score of between two to four indicated partial agreement; and one, strong disagreement.

² Item 16 is the overall score used as the dependent variable for statistical analyses. An overall score one or two indicated serious or extensive shortcomings. Potentially important but not serious shortcomings for the webpage attracted a score three or four. A score of five indicated minimal shortcomings.

2.3.2.3 SAM+CAM

Scoring using SAM+CAM was as described by Helitzer et al. (2009), involving a 22-item inventory divided into six categories: content, literacy demand, numeracy, graphics, layout/typography, and learning stimulation/motivation (see Table 2.4 next page). Items were scored 2=superior; 1=adequate; 0=not suitable; for a total of 44 points. This was expressed as a percentage. Non-applicable items were excluded from

percentage calculations.

Table 2.4 SAM+CAM inventory

Item	Scope	Score ¹
CONTENT		
1	Purpose	_____
2	Summary/review	_____
3	Desired reader behaviour	_____
4	Credibility	_____
LITERACY DEMAND		
5	Writing style	_____
6	Vocabulary helpers	_____
7	Confusion reducers	_____
8	Context	_____
9	Scope and length	_____
NUMERACY		
10	Numeric presentation	_____
11	Calculation	_____
GRAPHIC MATERIAL		
12	Document clarity	_____
13	Illustrations	_____
LAYOUT AND TYPOGRAPHY		
14	Layout and organisation	_____
15	Typography	_____
16	Subheading or advance organisers	_____
LEARNING STIMULATION AND MOTIVATION		
17	Motivators to attend to text	_____
18	Inclusion	_____
19	Reader interaction	_____
20	Theoretical application	_____
21	Tone	_____
22	Persuasive techniques	_____

¹ Scoring is out of three, with a score of zero indicating that the publication is not suitable, and a score of one being adequate. The highest score is two, indicating superior suitability.

2.3.2.4 Plain language

A custom plain language tool was adapted for this thesis from existing material based on: National Adult Literacy Agency (NALA) Checklist for Documents, and Plain Language Action and Information Network (PLAIN) Checklist for Plain Language on the Web (see Table 2.5).

The final checklist was the result of two Master of Audiology students, both involved in health literacy research (present author included), deciding on which NALA and PLAIN items to include. Consideration was given to ease of use, relevance to audiology themed websites, and clarity of item instruction. In cases where search engine results ranked videos within the top 10 hits (for given search terms), videos were transcribed using an online tool, www.diycaptions.com. Plain language analysis was then applied directly to transcriptions.

Table 2.5 Plain Language Checklist

Item	Description	Score
1	Does one or more of the headings contain the web search term?	_____
2	Does the introduction (first paragraph) inform the reader what they are about to read?	_____
3	Is the content relevant to the search terms used?	_____
	Does the material begin with the most important message?	_____
4	Is the content arranged in an order that makes sense?	_____
5	Are different topics grouped under separate headings or subheadings?	_____
6	Are personal pronouns such as “you” and “we” used throughout?	_____
7	Is an active voice used throughout?	_____
8	Are lay terms predominately used throughout?	_____
9	If technical terms are used, are they explained?	_____
10	Are simple sentences used throughout (i.e. no more than one new point per sentence)?	_____
11	Is correct grammar and punctuation used throughout?	_____
12	Are unnecessary words eliminated (e.g. technical jargon or adverbs)?	_____
13	Is the appearance of the material consistent throughout (i.e. consistent use of fonts, italics, bold print, colour, and bullet points)?	_____
14	Does the material look easy to read, with an uncluttered layout, plenty of white space, and dark coloured text on a light background or light coloured text on a dark background?	_____
15	Are the fonts clean in their design and easy to read (i.e., not cursive or uncommon)?	_____
16	Is the text size large enough for easy reading and does each line have about 10-15 words?	_____
17	Are Italics, underlining, capitalisation, and bold print used sparingly?	_____
18	Are images clear and uncluttered and related to the content?	_____

Adapted from (1) National Adult Literacy Agency: Plain English Checklist for Documents; (2) Center for Health Literacy: Quick Checklist for Plain Language; (3) Plain Language Action & Information Network: Checklist for Plain Language on the Web. Typeface, size, spacing, and contrast suggestions based on Rello and Baeza-Yates (2017). Items 1-3 assess webpage elements that facilitate reader focus. Items 5-6 assess organisation and layout. Items 7-13 assess writing. Items 14-19 assess design and formatting.

2.3.3 Reliability

All webpages assessed in this study were checked for HONcode certification. This was a simple check involving installation of the HONcode Toolbar for Mozilla Firefox (66.0.5; extension 3.1.3). Websites with current certificates (as of calendar year 2019) were scored “valid”

and those without, “expired or absent.”

2.3.4 Inter-rater reliability

This Master of Audiology thesis project complemented one of several related projects simultaneously run in the Kelly-Campbell lab. As such, prior to running a full analysis it was important to first establish internal consistency between raters’ overall scores. To achieve this individual rater variance structure was compared to total rater variance using coefficient alpha (Cronbach, 1951). This was undertaken in succession: (1) via pilot reliability analysis involving three practice webpages related to audiology but not included in the wider sample; (2) round one of reliability analysis consisting of $n=4$ randomised webpages sampled from within each researcher’s study and ($n=4$) randomised webpages sampled from the six other each researchers’ list of websites; and finally (3) round two of reliability analysis consisting of a second randomised set ($n=19$) of webpages drawn from all seven researchers’ lists of webpages. ICC was applied to raters’ overall PEMAT (understandability), PEMAT (actionability), Plain Language, and SAM+CAM¹ scores. The kappa generated from the ICC was used to determine how reliable the ratings were. Values of kappa range from 0 (no agreement between raters) to 1.0 (perfect agreement). According to Joseph L. Fleiss and Paik (2003), kappa values greater than 0.75 indicate “excellent agreement beyond chance” (Joseph L. Fleiss and Paik, 2003, p. 604). Values between .60 and .75 indicate good agreement beyond chance. Values between .40 and .59 indicate fair agreement beyond chance.

2.3.5 Statistical analyses

RStudio (Version 1.2.1335) was used for the following methods: Cohen’s D (reported as η^2) “lsr”, summary statistics (mean, (M), standard deviation (SD), median (Mdn), interquartile range (IQR), confidence interval (CI)), student’s t-test, mann-whitney test, intraclass correlation

¹Only two of seven researchers utilised the SAM+CAM inventory

(ICC), specifically the ICC(2,1) form as per (Shrout et al., 1979). R libraries used included: “irr,” “base,” and “psych.” JASP (Version 0.10.2) was used for all other methods: MANOVA, ANOVA, Kruskal-Wallis, correlational matrices (Spearman’s Rho), tables and figures

Chapter 3

Results

3.1 Survey response

Recruitment yielded 25 participant responses recorded between 27-May to 06-June 2019 (inclusive), all of whom met inclusion criteria of English language fluency. Twenty-three survey respondents provided three or more search terms. Table 3.1 summarises respondent gender, education, ethnicity and age demographics. Figure 3.1 shows the distribution of respondent ages.

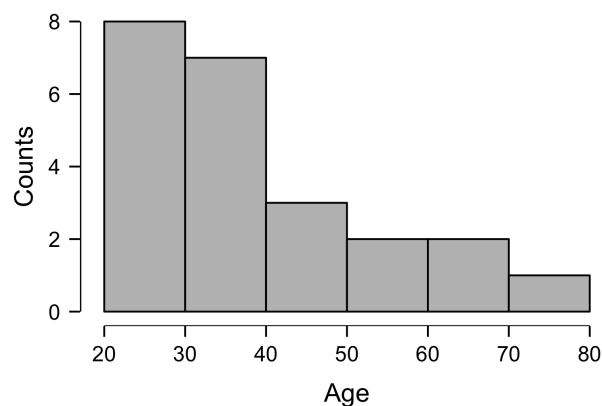


Figure 3.1 Counts of individual respondent age, bin size = 10 years

Table 3.1 Respondent gender, education, ethnicity

	<i>n</i>
GENDER	
Male	7
Female	16
Gender diverse	0
Unspecified	0
EDUCATION	
High school	2
Some tertiary-level	0
Bachelors	7
Masters/Doctorate	14
Unspecified	0
ETHNIC GROUP	
European	14
Māori	0
Pacific Peoples	1
Asian	3
Middle Eastern/Latin American/African	0
Other	5

Notes: Ethnic group “Other” included user-specified groups not elsewhere included (e.g., “Celtic”) in addition to multiple selections (e.g., “European” and “Asian”). Education group “Some tertiary-level” refers to study at university and polytechnic institutions.

Table 3.2 Respondent age, terms supplied, and time taken to complete survey

	Age	Search terms supplied	Time taken (minutes)
Mean	39.4	3.7	10.4
Std. Deviation	14.6	1.3	21.0
Minimum	22.0	1.0	1.4
Maximum	71.0	7.0	97.1

3.2 Search terms

In total, respondents provided 87 search terms. Distilling these down by similarity yielded 14 unique terms (e.g., “cochlea implant” was merged with “cochlear implant”) and common theme (e.g., “implantable hearing device,” and “hearing loss implants” were simplified to “hearing implant”). These are listed in Table 3.3. Ranking distilled terms according to the number of times they appeared in survey responses revealed that 86% of original search queries involved just five distilled search terms. Entering these top five terms in Google Trends revealed a number of related terms, as recorded by Google over the five-year period ending 06-July-2019. Taking into account those terms that were ranked, and trending terms not otherwise included in the distilled list, a final list of globally relevant (i.e., across English-speaking countries) search terms was generated (see Table 3.4).

Table 3.3 List of search terms refined from survey responses

Refined search term	<i>n</i>
Hearing implant ¹	56
Hearing device ¹	51
Hearing aid	12
Cost of hearing implant	7
Cochlear implant	6
Negative outcomes of hearing implant	7
Indications of hearing implants	3
Hearing loss	3
Surgery involved in hearing implantation	2
Deafness	2
Reliability of hearing implant	1
Manufacturers of hearing implants	1
Bone conduction hearing implants	1
Hearing implant technology	1

¹ Not necessarily mutually exclusive

Table 3.4 Refined search terms for Google Trends analysis

Input list ¹	Output list from Google Trends ²
Cochlear implant	Cochlear implant cost Cochlear implant MRI Cochlear implant surgery How does a cochlear implant enable the deaf to hear What is a cochlear implant Cochlear implant NZ
Hearing device	Hearing aid device BAHA implant
Hearing implant	Conductive hearing loss Sensorineural hearing loss

¹ Full list of refined search terms in Table 3.3 served as input terms for Google Trends analysis. However, only three of which generated sufficient data in Google Trends; these are shown here as “Input list.”

² Top 10 most relevant Google Trends terms concerning implantable hearing devices at time of writing. “Cochlear implant NZ” was unique to the google.co.nz domain and was included here because most survey participants resided in New Zealand at time survey taking. U.S., U.K., and Australia all shared in common cochlear-implant-related terms (verbatim or substantially similar) listed above.

3.3 Inter-rater agreement

Average score intraclass correlation (ICC) was used to measure the level of agreement between raters’ suitability scores. This was analysed using a one-way random effects model with suitability score (for each suitability inventory) as the dependent variable. The independent (random effects) variable in this case was Raters (for round one, seven levels: rater 1, rater 2 ... rater 7; and for round two, two levels: rater 1, rater 2). Under Fleiss’ six forms of ICC (Section 2.3.5), ICC(2,1) is the most appropriate ICC method for this data in relation to the question being asked. All reported ICC coefficients refer to ICC(2,1).

Round one coefficients for inter-rater agreement were as follows: PEMAT (understandability) ICC = .76; PEMAT (actionability) ICC = .42; plain language ICC = .69. Round two coefficients for inter-rater agreement were: PEMAT (understandability) ICC = .79; PEMAT (actionability) ICC = .97; plain language ICC = .70; DISCERN ICC = .80; SAM+CAM ICC = .089.

3.4 Readability of webpages

Ninety-eight webpages pages met inclusion criteria defined in Section 2.2.2. Mean readability scores and proportions at given RGLs for these are shown in Table 3.5. Spearman rank-order correlation revealed significant association between readability inventories (see Table 3.6).

Table 3.5 Reading grade level for all webpages, excluding videos

	FK	GF	SMOG
Mean	10.86	12.49	12.25
Std. Deviation	2.16	2.33	1.74
$\leq 6^{th}$ RGL	1%	0%	0%
$\leq 8^{th}$ RGL	5.1%	2%	1%
9^{th} RGL	14.3%	10%	6.1%
$10-12^{th}$ RGL	45.9%	50%	58/98
$>12^{th}$ RGL	25.5%	53.1%	59.2%

Table 3.6 Spearman Correlations

		GF	FK	SMOG
GF	Spearman's rho	—	—	—
	p-value	—	—	—
FK	Spearman's rho	0.94	—	—
	p-value	< .001	—	—
SMOG	Spearman's rho	0.95	0.95	—
	p-value	< .001	< .001	—

3.4.1 Location

Webpages classified by WHO geographic region numbered as follows: Americas ($n = 45$), Europe ($n = 11$), West Pacific ($n = 27$) and World/Other ($n = 15$). No webpages were sampled from Africa, South-East Asia or Eastern Mediterranean. As a result of low sample sizes among webpages for hosting organisations located in Europe and World/Other, these data were merged to improve balance ($n = 26$).

MANOVA testing revealed no statistically significant effect of location (Approx. $F_{(2,96)} = 1.2, p = .31$, Wilk's $\lambda = .93$, see Figure 3.2). Assumption checks using Box's M-test revealed heterogeneity of covariance ($\chi^2_{(30)} = 50.13, p = .0012$), and multivariate non-normality (Shapiro-Wilk = .88, $p < .001$). Marginal means for RGL by categorical variable WHO region ($k = 3$) are shown in Table 3.7.

Table 3.7 Marginal RGL means - WHO region

Location	Mean RGL	SE	95% CI	
			Lower	Upper
Americas	12.03	.35	11.34	12.72
W. Pacific	11.79	.38	11.04	12.55
World/Other	11.77	.38	11.02	12.52

Mean RGL refers to grand mean combining Flesch-Kincade reading grade level, Gunning-Fog index and simple measure of gobbledegook.

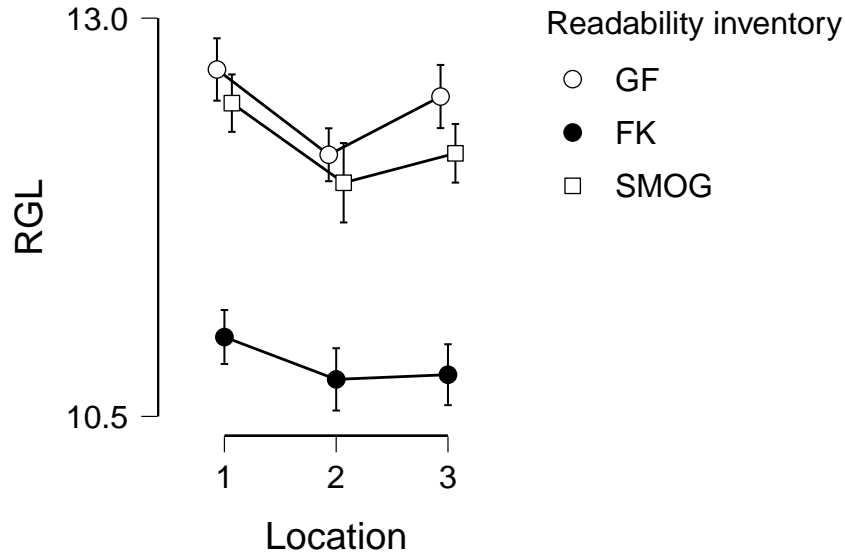


Figure 3.2 Mean RGL (\pm 95% CI) among webpages hosted by organisations located in the Americas (1), West Pacific (2), and World/Other (3). FK = Flesch-Kincaid, GF = Gunning Fog, SMOG = simple measure of gobbledegook.

3.4.2 Organisation type

Organisations were initially classified into one of five types. Two were not-for-profit (NFP) entities (government $n = 15$, non-government $n = 28$); two were for-profit (FP) entities (business-to-business $n = 40$, business-to-consumer $n = 12$); and a final category was blog sites and user forums ($n = 3$). To improve balance the latter was merged with NFP; and government/non-government subgroups were merged into one NFP group ($n = 46$). Similarly, business-to-business and business-to-consumer subgroups were merged in favour of a single group, FP ($n = 52$).

MANOVA testing revealed a significant main effect of organisation type (Approx. $F_{(1,97)} = 3.22, p = .0026$, Wilk's $\lambda = .90$; see Figure 3.3.). See Section 3.4.1 for assumption checks using Box's M-test multivariate normality. Grand mean RGL by categorical variable organisation type ($k = 2$) are shown in Table 3.8. Follow-up two-way ANOVA with fixed factors organisation type (two levels: FP, NFP) and readability inventory (three levels: Gunning-Fog, Flesch-Kincaid, SMOG) revealed no

significant difference between FP ($M = 11.76$, $SD = 2.042$) and NFP ($M = 11.98$, $SD = 2.38$) $F_{(1,288)} = .815, p = .37, \eta^2 = .003$.

Table 3.8 Marginal RGL Means - Organisation Type

Organisation type	Marginal Mean	SE	95% CI	
			Lower	Upper
FP	11.76	0.28	11.19	12.33
NFP	11.98	0.302	11.38	12.58

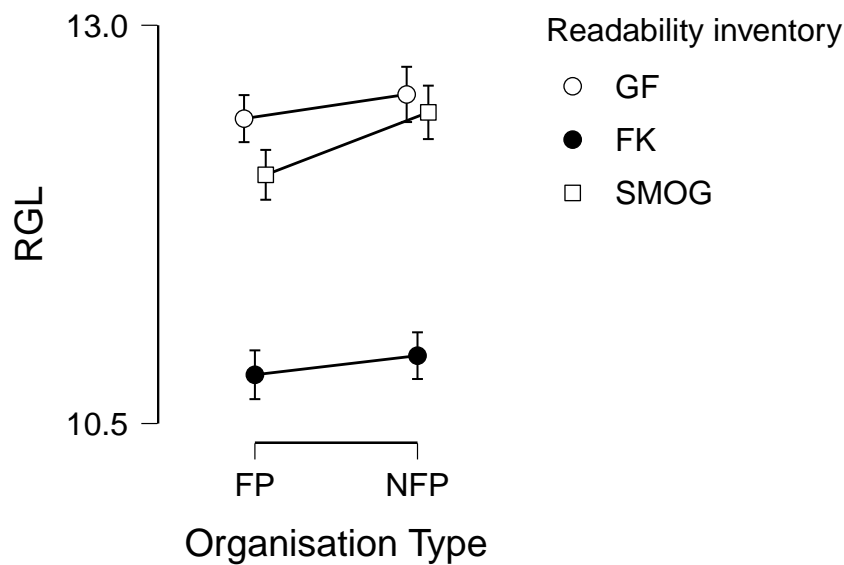


Figure 3.3 Mean RGL scores (\pm 95% CI) among webpages hosted by for-profit (FP) and not-for-profit (NFP) organisations. FK = Flesch-Kincaid, GF = Gunning Fog, SMOG = simple measure of gobbledegook.

3.4.3 Interaction between organisation type and location

MANOVA testing revealed no significant interaction between location and organisation type (Approx. $F_{(2,96)} = .82, p = .55$, Wilk's $\lambda = .95$. See Section 3.4.1 for assumption checks using Box's M-test multivariate normality. Marginal means are provided in Table 3.9.

Table 3.9 Marginal RGL means - Location x Organisation Type

Location	Organisation Type	Mean RGL	SE	95% CI	
				Lower	Upper
Americas	FP	12.19	.26	11.68	12.71
	NFP	11.92	.25	11.43	12.41
W. Pacific	FP	11.04	.29	10.46	11.62
	NFP	12.59	.38	11.84	13.34
World/Other	FP	11.99	.32	11.35	12.62
	NFP	11.59	.35	10.91	12.29

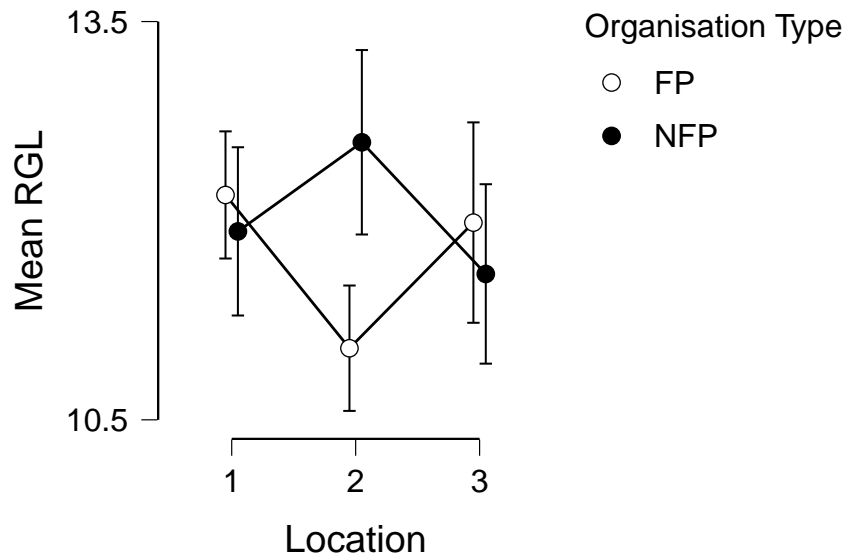


Figure 3.4 Mean RGL among webpages with nesting on organisation type (for-profit, FP; not-for-profit, NFP) and WHO region (Americas, 1; West Pacific, 2; World/Other, 3) Error bars = 95% CI Note: Mean RGL refers to mean across GF, FK, and SMOG; i.e., consistent with MANOVA analysis.

3.4.4 HONcode

Readability data was grouped according to those webpages that contained a valid HONcode certificate (“Yes” $n = 13$), and compared to those without (“No” $n = 85$). To examine potential differences in readability formula sensitivity, FK, GF, and SMOG readability scores were analysed separately using t-tests with Bonferroni correction ($\alpha = .05/3 = .017$).

FK readability data was homoskedastic (Bartlett’s K-squared₍₁₎ = .33, $p = .56$) and normally distributed (Shapiro-Wilk: “Yes” $W = .95, p = .65$, “No” $W = .98, p = .22$). An independent two-sample Student’s t-test revealed no significant difference in FK readability scores between websites with HONcode certification ($M = 10.92, SD = 1.93$) compared to those without ($M = 10.85, SD = 2.2$), $t_{(96)} = .97, p = .92, \eta^2 = .029$.

GF readability data was homoskedastic (Bartlett’s K-squared₍₁₎ = .056, $p = .81$) and normally distributed (Shapiro-Wilk: “Yes” $W = .96, p = .69$, “No” $W = .98, p = .32$). An independent two-sample Student’s t-test revealed no significant difference in GF readability scores between websites with HONcode certification ($M = 12.7, SD = 2.24$) compared to those without ($M = 12.45, SD = 2.4$), $t_{(96)} = .35, p = .73, \eta^2 = .11$.

SMOG readability data was homoskedastic (Bartlett’s K-squared₍₁₎ = .0014, $p = .97$) and normally distributed (Shapiro-Wilk: “Yes” $W = .98, p = .98$, “No” $W = .99, p = .86$). An independent two-sample Student’s t-test revealed no significant difference in SMOG readability scores between websites with HONcode certification ($M = 12.4, SD = 1.74$) compared to those without ($M = 12.22, SD = 1.75$), $t_{(96)} = .34, p = .73, \eta^2 = .11$.

3.5 Suitability of webpages

Suitability scores for all websites ($n = 98$) are shown in Table 3.10. Spearman rank-order correlation revealed significant association between suitability inventories (see Table 3.11).

Table 3.10 Suitability scores for all webpages

	PEMAT		Plain Language	DISCERN	SAM+CAM
	U	A			
Mean	69.76	40.39	73.67	2.91	43.73
SD	15.34	33.49	16.06	.98	14.72

PEMAT understandability and actionability indicated by U and A, respectively. DISCERN score refers to item 16 “overall rating of publication,” see Table 2.3 in Section 2.3.

Table 3.11 Spearman correlations between suitability inventories

		P-U	P-A	PL	D	S+C
P-U	Spearman's rho	–				
	p-value	–				
P-A	Spearman's rho	0.463	–			
	p-value	< .001	–			
PL	Spearman's rho	0.73	0.59	–		
	p-value	< .001	< .001	–		
D	Spearman's rho	0.55	0.42	0.56	–	
	p-value	< .001	< .001	< .001	–	
S+C	Spearman's rho	0.68	0.59	0.76	0.59	–
	p-value	< .001	< .001	< .001	< .001	–

3.5.1 Location

DISCERN and PEMAT (actionability) scores, by location type, exhibited non-normal, non-continuous, and platykurtic distributions. Therefore, a rank-based one-way ANOVA (Kruskal-Wallis) with Bonferroni correction ($\alpha = .05/3 = .017$) was applied separately to DISCERN and PEMAT (actionability) data, grouped by the independent variable WHO location (three levels: Americas, W. Pacific, World/Other). Note: median (*Mdn*) and interquartile range (*IQR*) are reported in place of mean (*M*) and standard deviation (*SD*), i.e., for consistency with non-parametric ANOVA testing.

No significant difference was found among PEMAT (actionability) scores for webpages located in the Americas (*Mdn* = 50, *IQR* = 80), W. Pacific (*Mdn* = 40, *IQR* = 80), or World/Other (*Mdn* = 40, *IQR* = 60) Kruskal-Wallis = .746₍₂₎, *p* = .69.

Similarly, no significant difference was found among DISCERN scores for webpages located in the Americas (*Mdn* = 3, *IQR* = 1), W. Pacific (*Mdn* = 3, *IQR* = 2), or World/Other (*Mdn* = 3.18, *IQR* = 1.8) Kruskal-Wallis = 1.07₍₂₎, *p* = .59.

SAM+CAM, PL, and PEMAT (understandability) MANOVA testing revealed no statistically significant main effect of WHO location (Approx.

$F_{(2,95)} = .69, p = .65$, Wilk's $\lambda = .96$; see Figure 3.5). Assumption checks using Box's M-test revealed homogeneity of covariance ($\chi^2_{(30)} = 35.35, p = .23$), and multivariate non-normality (Shapiro-Wilk = .97, $p < .01$). Marginal means for this data is shown in Table 3.12.

Table 3.12 Marginal means (%) - Suitability by WHO region (SAM+CAM and PEMAT (understandability))

Location	Mean (%)	SE	95% CI	
			Lower	Upper
Americas	64.05	2.35	59.38	68.72
W. Pacific	61.59	2.55	56.53	66.66
World/Other	61.52	2.57	56.42	66.62

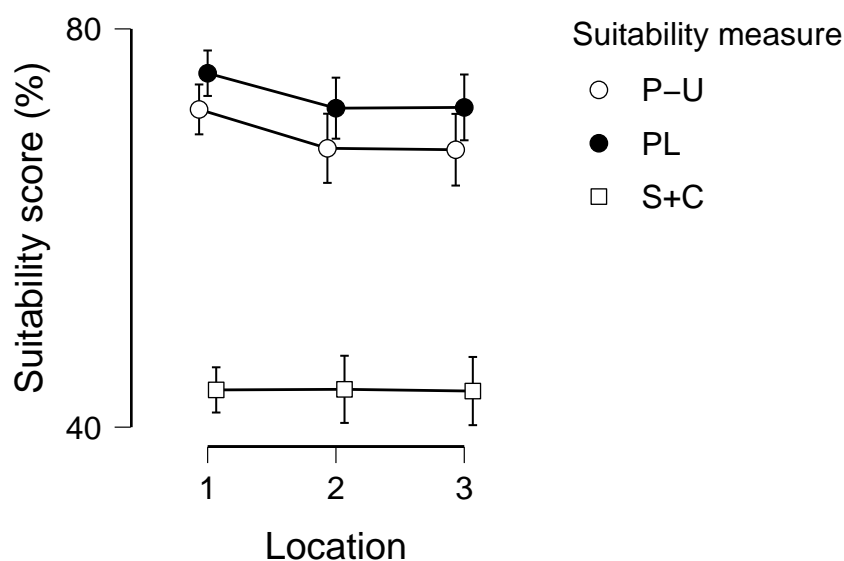


Figure 3.5 Suitability scores for PEMAT understandability (P-U), plain language (PL), and SAM+CAM (S+C) for webpages hosted by organisations located in the Americas (1), West Pacific (2), and World/Other (3). Error bars = 95% CI

3.5.2 Organisation type

As above in Section 3.5.1, PEMAT (actionability) and DISCERN suitability scores were analysed separately using a rank-based one-way

ANOVA (Kruskal-Wallis). These were grouped by the independent variable organisation type (two levels: FP, NFP). Bonferroni correction ($\alpha = .05/3 = .017$) was applied.

No significant difference was found in PEMAT (actionability) scores of webpages of for-profit entities ($Mdn = 50, IQR = 80$) compared to not-for-profit entities ($Mdn = 40, IQR = 60$) Kruskal-Wallis = $3.13_{(2)}, p = .077$.

Similarly, no significant difference was found in DISCERN scores for webpages of for-profit entities ($Mdn = 3, IQR = 1$) compared to not-for-profit entities ($Mdn = 3, IQR = 1$) Kruskal-Wallis = $2.67_{(2)}, p = .102$.

MANOVA (as described above in Section 3.5.1) revealed no significant effect of organisation type (Approx. $F_{(1,96)} = 1.16, p = .33$, Wilk's $\lambda = .96$, see Figure 3.6). Marginal means for this data is shown in Table 3.13.

Table 3.13 Marginal means (%) - Suitability by organisation type (excluding PEMAT (actionability) and DISCERN)

Suitability measure	Marginal Mean	SE	95% CI	
			Lower	Upper
P-U	69.65	1.57	66.54	72.76
PL	73.44	1.57	70.33	76.55
S+C	44.074	1.57	40.96	47.19

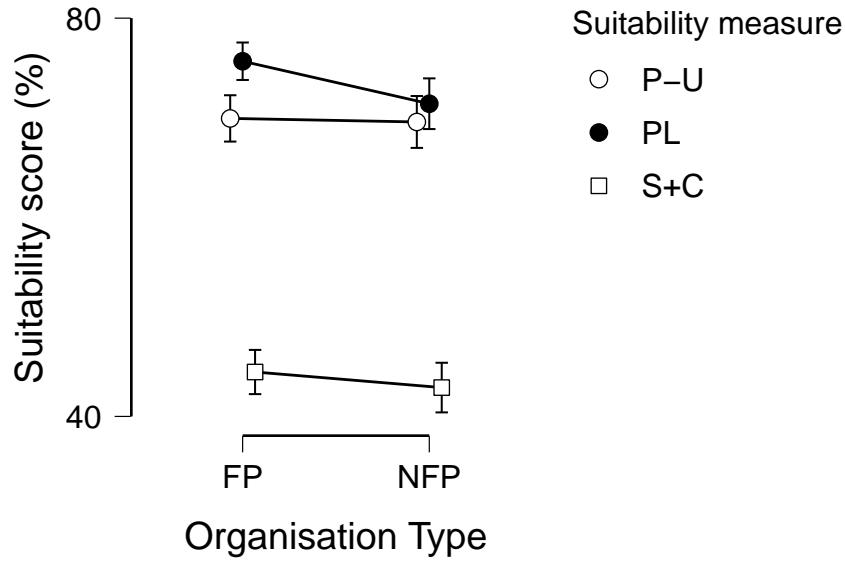


Figure 3.6 Suitability scores for PEMAT understandability (P-U), plain language (PL), and SAM+CAM (S+C) for webpages hosted by for-profit (FP) and not-for-profit (NFP) organisations. Error bars = 95% CI

3.5.3 Interaction between organisation type and location

The same MANOVA model used for testing the above main effects of WHO region (across all levels of suitability, irrespective of organisation type), and organisation type (across all levels of suitability, irrespective of WHO region), was also used to test the interaction between these two categorical variables. No statistically significant interaction was found (Approx. $F_{(2,95)} = 1.49, p = .18$, Wilk's $\lambda = .91$, see Figure 3.7).

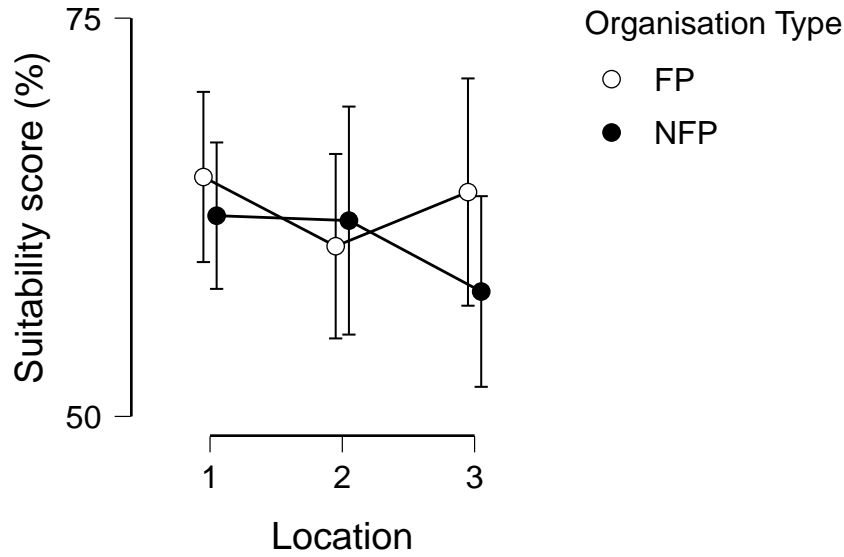


Figure 3.7 Interaction between WHO region and type of organisation for web page suitability. Mean RGL among webpages with nesting on organisation type (for-profit, FP; not-for-profit, NFP) and WHO region (Americas, 1; West Pacific, 2; World/Other, 3) Error bars = 95% C.I. Note: Mean suitability score (%) refers to mean across PEMAT (understandability), plain language, and SAM+CAM; i.e., consistent with MANOVA analysis.

3.6 Reliability: HONcode

Suitability data was grouped according to those webpages that contained a valid HONcode certificate (“Yes” $n=13$), and compared to those without (“No” $n=85$). The procedures used for this analysis were similar to those used for readability in Section 3.4.4.

Suitability inventories (grouped by the independent variable) were analysed separately because of differences in data types. Discrete interval data included PEMAT actionability, and plain language. Discrete ordinal data included DISCERN. Continuous data included PEMAT understandability, and SAM+CAM. In order to compare the effect of HONcode certification on suitability scores, data were analysed (where appropriate) using parametric or non-parametric methods. Bonferroni

correction was applied ($\alpha = .05/5 = .01$).

PEMAT actionability data followed non-normal distributions (Shapiro-Wilk: “Yes” $W = .78, p = .0041$, “No” $W = .86, p < .001$), with equal variance between levels of the independent variable (Bartlett’s $K\text{-squared}_{(1)} = .34, p = .56$). An independent two-sample Mann-Whitney test with correction for ties revealed no significant difference in PEMAT understandability scores between websites with HONcode certification ($Mdn = 0, IQR = 40$) compared to those without ($Mdn = 50, IQR = 80$), $W = 367, p = .046$.

Plain language data in the case of HONcode certified webpages was normally distributed (Shapiro-Wilk: $W = .94, p = .44$), and non-normally distributed among those without HONcode certification (Shapiro-Wilk: $W = .97, p = .042$). Equal variance was found between levels of the independent variable (Bartlett’s $K\text{-squared}_{(1)} = .65, p = .42$). An independent two-sample Mann-Whitney test with correction for ties revealed no significant difference in plain language scores between websites with HONcode certification ($Mdn = 80, IQR = 20$) compared to those without ($Mdn = 80, IQR = 25$), $W = 581.5, p = .76$.

DISCERN data followed non-normal distributions (Shapiro-Wilk: “Yes” $W = .67, p < .001$, “No” $W = .91, p < .001$), with equal variance between levels of the independent variable (Bartlett’s $K\text{-squared}_{(1)} = 7.69, p = .0056$). An independent two-sample Mann-Whitney test with correction for ties revealed no significant difference in DISCERN scores between websites with HONcode certification ($Mdn = 3, IQR = 0$) compared to those without ($Mdn = 3, IQR = 2$), $W = 606, p = .56$.

PEMAT understandability data followed non-normal distributions (Shapiro-Wilk: “Yes” $W = .83, p = .016$, “No” $W = .97, p = .042$), with equal variance between levels of the independent variable (Bartlett’s $K\text{-squared}_{(1)} = .12, p = .73$). An independent two-sample Mann-Whitney test with correction for ties revealed no significant difference in PEMAT understandability scores between websites with HONcode certification ($Mdn = 75, IQR = 9.8$) compared to those without ($Mdn = 69, IQR = 19.8$), $W = 558, p = .96$.

SAM+CAM data followed normal distributions (Shapiro-Wilk: “Yes” $W = .96, p = .81$, “No” $W = .97, p = .096$), with equal variance between levels of the independent variable (Bartlett’s $K\text{-squared}_{(1)} = 1.96, p = .16$). An independent two-sample Student’s t -test revealed no significant difference in SAM+CAM scores between websites with HONcode certification ($M = 40.82, SD = 10.87$) compared to those without ($M = 44.17, SD = 15.22$), $t_{(96)} = .76, p = .45, \eta^2 = .23$.

3.7 Videos

Search results included a total of 11 videos, two of which contained no dialogue or on-screen text, only animations with background music. These two videos were excluded from readability and PEMAT analysis.

YouTube hosted eight of nine videos, Khan Academy the other; neither of which contained HONcode certification.

Transcriptions of video dialogue to text files on average were $Mdn = 426$ words per transcription ($IQR = 294$). This consisted of FP entities ($n = 8$) and a single NFP entity; all which were located in the WHO location, Americas. Thus, no readability and suitability (PEMAT) comparisons could be made among video transcription data because both independent variables (location, organisation type) contained one level with zero variance. Instead, a comparison was made between transcriptions and webpages. For consistency, webpage data was limited to FP entities, without HONcode certification, located in the Americas. Comparisons were made using five separate Mann-Whitney tests. Bonferroni correction was applied ($\alpha = 0.05/5 = 0.01$).

GF scores for video transcriptions ($Mdn = 12.21, IQR = 4.09$) compared to webpages ($Mdn = 13.29, IQR = 2.45$) were not significantly different $W = 69, p = 0.711$. SMOG scores for video transcriptions ($Mdn = 12.10, IQR = 2.23$) compared to webpages ($Mdn = 13.05, IQR = 1.66$) were not significantly different $W = 60.5, p = 0.403$ (continuity correction for ties). FK scores for video transcriptions ($Mdn = 10.07, IQR = 3.78$) compared to webpages ($Mdn = 11.54, IQR = 3.34$) were not significantly different

$W = 61, p = 0.43$. PEMAT (understandability) scores for video transcriptions ($Mdn = 44, IQR = 12.5$) compared to webpages ($Mdn = 70.6, IQR = 22.3$) were significantly different $W = 11.5, p < 0.001$ (continuity correction for ties). PEMAT (actionability) scores for video transcriptions ($Mdn = 0, IQR = 80$) compared to webpages ($Mdn = 66.7, IQR = 40$) were not significantly different $W = 57, p = 0.29$ (continuity correction for ties).

Chapter 4

Discussion

The purpose of this thesis was to investigate how readable and how suitable webpages were for information seekers wishing to learn more about the general topic of implantable hearing devices. This included top-ranking webpages related to cochlear implants, bone-anchored hearing aids, SoundBridge devices, and auditory brainstem implants. To this end, webpages were assessed using widely published readability and suitability tools; the former: Gunning Fog, Simple Measure of Gobbledygook, and Flesch-Kincaid; the latter: Plain Language, DISCERN, SAM+CAM, and PEMAT. Our goal was answer the following questions: firstly, did readability and suitability scores differ depending on global locality of English webpages; secondly, did organisation type of each domain matter in terms of readability and suitability scores; and thirdly, did webpages with HONcode certification differ in their readability and suitability scores compared to those without.

4.1 Webpage readability

Present findings show that internet information related to implantable hearing devices requires a reading grade level consistent with that of a 15 to 17 year-old high school student (RGL 10-12). For New Zealand students under the NCEA system this would correspond to years 11

through 13.

As shown in Table 3.5, one webpage out of 98 in this study scored less than a 6th RGL (although only with the FK formula). This particular finding was unfortunate in light of the American Medical Association, and Institute for Healthcare Advancement both having issued recommendations—over a decade ago—that patient health education materials not exceed a 6th RGL (Weiss, 2003; Mayer and Villaire, 2007). Interestingly, these recommendations were pitched toward clinicians, and presumably medical writers. That fact, however, does not anticipate very well present findings: for example, Stanford Children’s Health, Johns Hopkins Medicine, and WebMD, all scored RGLs within the range 9–11; which does not even meet the National Center for Education Statistics’ recommendation of 8th RGL or less for educational print materials (NCES, 2019). In fact, only 5% of webpages sampled in the present thesis satisfied the latter recommendation.

That such a low proportion of websites met this criteria was not altogether surprising in light of previous work by others. For instance, Boston et al. (2004) found of the 42 cochlear implant websites included in their study, a 12th RGL was required for comprehension. Both SMOG and GF inventories used in the present study required 12th RGL. As shown in Table 3.5 however the FK method was indicative of a slightly lower RGL, which was interesting because Begeny and Greene (2014) showed that SMOG and FOG both showed bias toward the lower RGLs (although only grades 1 through 7 were considered)(Begeny and Greene, 2014). Notwithstanding, all three readability inventories in the present study were highly cross-correlated, further suggesting same construct (see Table 3.6). These trends are concerning given the apparent confidence some New Zealanders (and presumably other populations) with hearing impairment have finding online health information (Peddie and Kelly-Campbell, 2017).

A handful of implantable hearing device related studies report comparable findings. For instance, Boston et al. (2004) carried out a survey aimed at parents of children who completed surgery for cochlear implantation. They found that parents’ preferred source of information

remained health care professionals, despite having accessed internet information on cochlear implant brand choice, surgery details and communication portals (forums) with other parents. Boston et al. (2004) also found C.I. information on the internet to be incomplete, misleading, requiring a 12th grade reading level according to the Fry Readability Graph (Fry, 1979), which the authors found troubling given that 80% of parents surveyed used the internet for gathering cochlear implant information prior to their child's surgery (in addition to seeking advice from health care professionals).

As with the approach taken in this thesis, Seymour et al. (2015b) assessed "cochlear implant" and "cochlear implant surgery" webpages' for their readability. Unlike present methodology, Seymour et al. (2015b) included only those websites written for the general public by a medical professional. Scientific and medical websites, those of biotech companies, adverts, blogs, and newspapers were all excluded. However, in agreement with present findings, the authors found that the average reading grade level was 10-12 using the Flesch-Kincaid method, and 13 for Gunning-Fog (cf. Table 3.5).

Other recent studies have found patient education materials too difficult to read for most parents on topics related to newborn hearing screening, microtia, atresia, tympanostomy, and tinnitus (Sax et al., 2019; Harris et al., 2018; Alamoudi and Hong, 2015b; Manchaiah et al., 2019). Similar trends are reported for medical topics unrelated to hearing (Doruk et al., 2018; Alsoghier et al., 2018; Castro-Sánchez et al., 2015; Doruk et al., 2018; Jo et al., 2018; Narwani et al., 2016; Wiriyakijja et al., 2016; Ruble et al., 2019).

Web materials falling within this range pose challenges for many adults. This is because half or more of adult populations among developed countries including New Zealand, Australia, the U.S., and, to a lesser extent, Canada—read at low literacy levels (Kaestle et al., 2001; Carroll, 1987; NCES, 2019; MoE, 2008).

Gaps between adult literacy and readability of online technical material demonstrates one potential way in which prose may be improved, namely

though use of shorter sentences consisting of words familiar to most adults. However, readability formulas alone are only part of the solution; and in some cases are part of the problem. Technical sentences shortened through use of acronyms, for example, may score well in terms of RGL yet register indecipherable even amongst health-care professionals.

Consider: “SmartSNR eQ with SEEK and ActiveFocus are at the core of all Neural XII Sound Processors.” This hypothetical sentence contains 16 words of 1.5 syllables on average per word, with a RGL of 8-11 (e.g., Flesch-Kincaid, SMOG). Accordingly, 13- to 16-year old students should find this readable, in principle.

Because of limitations in relying exclusively on readability formulas, limitations such as insensitivity to word order or grammatical complexity, misused words, acronyms, abbreviations, persuasive effectiveness, imagery, or style and organisation (Redish, 1981, 2000). One solution applied herein was to assess web materials using suitability formulas—in addition to readability formulas. Our suitability inventories included PEMAT, Plain language, DISCERN, and SAM+CAM.

4.2 Webpage suitability

Webpages assessed using the Plain language tool scored on average just over 70%, with Design and Formatting scoring the highest overall compared to Reader Focus (76%), and Organisation (78%). Writing style scored lowest (60%) overall within sampled implantable device webpages, in large part due to more than four out of five websites using passive tense. On the other hand grammar and punctuation tended toward near perfect accuracy, while approximately half of all websites used long and complicated sentences (i.e., containing more than one idea) without sufficient explanation of technical terms. Webpages generally made use headings and subheadings followed by relevant content, aiding reader focus; however, only half of webpages informed the reader of what they were about to read in the introductory paragraph. Added to this, a similar proportion of webpages contained cluttered or irrelevant imagery,

either in-line or to the side of main text in the form of self-promotion and other forms of advertising. As a result, information seekers require deeper scrolling into pages to determine whether or not the content bears relevance to their initial search engine term(s). Together with insufficient technical explanation of medical and audiological terms, long sentences, and potentially distracting imagery, it is expected that most adult readers would find implantable hearing device content somewhat challenging.

Consistent with the Plain Language tool, the PEMAT tool also revealed high overall webpage relevance in relation to the main topic at hand, with as little as 10% unnecessary distractions (item #1, Table 2.1). It should be pointed out that the way in which PEMAT and Plain Language establish relevance differ slightly. Whereas the latter specifies three items, PEMAT specifies only a single item for relevance without making a distinction as to where exactly on the webpage purpose is established. Nevertheless, PEMAT and Plain Language were highly correlated overall (see Table 3.11), which is to some extent expected given that they both measure similar constructs, e.g., explanations for technical terms, active vs. passive voice.

PEMAT also includes constructs that specifically assess how well each webpage prompts readers to take action (item #18 through 23, Table 2.1). We found that less than half of all webpages assessed identified a call-to-action readers could take; and of these, less than one-third broke these into manageable explicit steps. Particularly deficient were use of visual aids and explanations for visual aids which might otherwise help readers take action. Therefore, readers arriving at top-ranking webpages related to “cochlear implants,” for example, would for the most part be required to return to their search engine’s top hits in order to continue research. Such discontinuities may introduce further challenges with varied readability and suitability. Medical topics similarly assessed using the PEMAT tool show consistency with the present findings, with understandability scores of between 50 to 70% and actionability scores ranging from 20 to 50% (Kirby et al., 2018; Ruble et al., 2019; Doruk et al., 2018)

The third tool used, SAM+CAM, revealed similar overall literacy demand

compared with PEMAT and Plain Language. This was because just under half of webpages assessed scored the highest value “2” for use of active, personal, easy to understand writing style. Infrequent use of common, explicit, non-technical terms (i.e., one-third of webpages scored “2” for this inventory item) was a second reason literacy demand was somewhat high. A third reason was that half of webpages provided no explanations or analogies or other techniques for reducing confusion of technical terms and concepts. Further to PEMAT revealing poor use calls-to-action, SAM+CAM revealed that half the webpages assessed made no attempt at framing key messages to promote self efficacy. Hence, together with the point made above (readers necessarily having to return to search engine results to continue their research), not only would readers miss out on take home messages, confusion may actually be worsened by other websites omitting this important writing technique. Where webpages did score high with the SAM+CAM tool was with the overall layout; including frequent use of headings, bullets, easy to read (and appropriately sized) typefaces, and logical progression of ideas.

The fourth suitability tool employed was DISCERN, which, unlike the aforementioned suitability tools, included inventory items for assessing treatment choices. Here, extensive shortcomings were found with respect to outlining what would happen if no treatment was taken up, or treatment was postponed. Such risks are beyond the scope of this thesis to discuss in detail; however, it is worth to mention the importance of language development and communicative function in a psychosocial context. It is well known that deaf and hard of hearing children are vulnerable to physical and emotional abuse, even sexual abuse—more so than normal hearing children. Such adverse childhood experiences, some authors reason, forms the basis for elevated lifetime prevalence of psychiatric morbidities such as depression, internalising disorder, and oppositional defiant disorder (Glickman, 2013; Dreyzehner and Goldberg, 2019; Kvam et al., 2007; Fellingner et al., 2009; Van Gent et al., 2007). Other areas identified with serious shortcomings were lack of transparency concerning sources of information, and, by extension, when such information had been published. As a result, almost half of webpages

assessed made no attempt, or very poor attempts, at identifying areas of uncertainty and proving follow-up sources of information.

Similar to the present thesis, Seymour et al. (2015b) assessed “cochlear implant” and “cochlear implant surgery” webpages using the DISCERN tool and reported an average score of 45%. Since the present author did not compute an overall average across all items in the DISCERN inventory, for the purposes of comparison, present data were reanalysed¹ This gave a mean score of 52%, which is slightly higher although consistent with the 2015 study of Seymour et al. (2015b). Elsewhere, studies report comparable suitability results for otolaryngology- and audiology-related webpages, e.g., (Joseph et al., 2016; Harris et al., 2018; Ritchie et al., 2016; Seymour et al., 2015a; Manchaiah et al., 2019).

4.3 Global differences in webpage readability and suitability

Similarities in health literacy have been shown to exist across some English-speaking countries (MoE, 2008). Because of this, one of the aims of the present thesis was to investigate whether or not readability and suitability scores differed across global English-speaking regions. Few systematic reviews have looked into this in an audiological context. One such study by Laplante-Lévesque and Thorén (2015) revealed that across several original peer-review studies, RGLs of between 11–13 were typical for U.K. and international based websites, with U.S. based websites slightly lower at 9–11.

To extend this work WHO regional coding was applied to webpages based on locations of presumed target audiences. For instance, of those webpages relating to “implantable hearing devices” that fell under the auspices of Wikipedia, whose audience presumably stretches globally,

¹ $DISCERNscore(\%) = \frac{\sum_{i=1}^{16}}{16 \times 5} \times 100$ where,

numerator sum refers to items 1-16 on Table 2.3 and denominator refers to the maximum possible score across all 16 items.

coding was therefore designated “World/Other.” Examples of domains with more granular regional coding were New Zealand’s Bay Audiology, and Johns Hopkins in the U.S., herein respectively as “West Pacific” and “Americas.”

As shown in Fig. 3.2, an apparent trend emerged toward lower readability scores among Australian and New Zealand websites. However, multivariate methods failed to show any statistically significant difference between regions. As a result, no strong evidence was found to suggest that webpages related to implantable hearing devices differed in their readability scores between the Americas, West Pacific, and World/Other. Similarly, no evidence was found to suggest that webpages written for global audiences, the Americas, or those in West Pacific regions differed in their suitability scores. To the best of this author’s knowledge no previous work has been carried out with a similar aim, within the field of medical audiology. That said, previous research within wider medical contexts have been carried out with similar goals to ours, such as the study by Castro-Sánchez et al. (2015). In their study, Canadian, Australian and European based domains reportedly had 11th and upward RGLs using FK, SMOG, and GF readability scores; whereas, U.S. as well as global (i.e., World Health Organisation) based domains scored 8th to 11th RGLs. Robins et al. (2016) found, when assessing male fertility clinic website information, that RGLs from the Americas and international domains all exceeded 12th RGL

4.4 Commerical webpages

In answering the second question posed in this thesis, as to whether or not organisation type influenced readability and suitability amongst for-profit and not-for-profit entities, domains within our study sample were separated accordingly.

First, it was revealed that a 12th grade RGL was required for webpages of both entity types (see Table 3.8). As a result, no statistically significant difference was found in RGL. Second, concerning webpage suitability,

overall DISCERN ratings (item 16, see Table 2.3) were suggestive of potentially important but not serious shortcomings for either entity type. For this comparison, no statistically significant difference emerged.

These findings are consistent with other studies in the context of otitis media with effusion, and tinnitus; for instance, Ritchie et al. (2016); Manchaiah et al. (2019) both found no difference between organisation type (i.e., for-profit, not-for-profit) in terms of readability and suitability obtained using Flesch-Kincaid reading ease and DISCERN. The latter two scores were also consistent with those reported here in that audiological-related webpages scored a RGL of 10-12. Ritchie et al. (2016) also reported a DISCERN score of 55-60%. We reported a DISCERN score of 3 in Chapter 3 and a recalculated score of 52%¹. Admittedly their study concerned webpages related to otitis media with effusion, whereas ours concerned implantable hearing devices; however, both topics are interrelated to some extent being within the context of audiological management. Joseph et al. (2016) also report no difference in readability using the SMOG formula between organisation types, which overlapped with a number of for-profit entities found in this study. A related study concerning “glue ear” webpages similarly found an average FK reading ease of 50 (11thRGL), and DISCERN score of 57 (Ritchie et al., 2016). As with present findings and that of other studies, differences between for-profit and not-for-profit entities failed to reach statistical significance Ritchie et al. (2016); Joseph et al. (2016); Ritchie et al. (2016); Robins et al. (2016).

4.5 HONcode

The third and final question asked in this thesis was whether or not readability and suitability measures differed between those webpages with HONcode certification compared to those without. Statistical analyses failed to reveal any difference in readability and suitability for any inventory. Therefore, HONcode certification did not seem to improve (or worsen) how well information-seekers are able to read and understand

web materials related to implantable hearing devices. Nor was any evidence found in support of HONcode certification having an effect on relevance, page layout, numerical skills, and all the other constructs tested under the present suitability analyses.

To the best of this author’s knowledge no prior readability and suitability studies have compared English websites with HONcode certification to those without, in the context of implantable hearing devices. In other audiological contexts, Alamoudi and Hong (2015b) report that, of the 16 websites included in their study related to microtia none were HONcode certified, nor were webpages related to atresia with the exception of one. For tinnitus related webpages, Manchaiah et al. (2019) found 18 out of 134 sampled websites were HONcode certified. Similarly, Greer (2019) reported only three out of 44 assessed webpages related to tinnitus contained HONcode certification—although this study was specific to websites written in the Spanish language. A number of other studies unrelated to audiology have asked similar questions. For instance, Grewal and Alagaratnam (2013b) report no difference in DISCERN or RGL (FK) for colorectal disease websites with HONcode certification compared to those without. Other studies report comparably low numbers of HONcode certification among medical-information websites (Wiriyakijja et al., 2016; Alsoghier et al., 2018; Narwani et al., 2016; Jo et al., 2018; Kortlever et al., 2019).

Further to these findings, it is interesting to note that some parents of deaf children do not appear to be selecting sources of information at all based on HONcode certification. As one study shows, even Government’s websites—which happens to be HONcode certified—are completely ignored in some cases (Porter and Edirippulige, 2007).

4.6 Audiovisual materials

For reasons explained further in Section 4.7, video content and respective transcriptions (Section 2.3.2.4) were compared to that of non-audiovisual webpages. The latter proved significantly more understandable according

to the PEMAT tool than videos at 71% compared to 44%. Unexplained medical terminology was a common finding among videos, suggesting greater familiarity with medical concepts compared to non-audiovisual webpages. However, being that the sample size was small for videos related to implantable hearing devices (via search engine hits), it is difficult to read too much into this finding. Another reason video understandability rated lower was because content was often not broken down into shorter sections (i.e., single topic leading in to another), and thus helping to minimise information overload (Shoemaker et al., 2013c).

By comparison, A recent study by Kang and Lee (2019) examined diabetes audiovisual materials using the PEMAT tool. Although this study concerned materials in the Korean language, the authors report an average understandability score of 50%, which was consistent with 44% found here for implantable hearing devices. For actionability the authors report 31%, whereas reported here was 0%. Other medical studies report similar understandability scores for English language YouTube videos, suggesting that many patient education audiovisual materials are inappropriate for viewers with low health literacy (Salama et al., 2019; Bellon-Harn et al., 2020).

4.7 Limitations

Before conclusions can be drawn from this data, several aspects of the study deserve mention.

First of which, regional suitability and readability comparisons lacked granularity due to domain location imbalance. As a result, it is difficult to draw conclusions as to whether and to what extent European websites (i.e., bilingual including English) may differ to their U.S. counterparts, in terms of readability and suitability within the field of implantable hearing devices. In retrospect, such a limitation was probably inevitable; Google ranks webpages in part by keyword density, which corporate for-profit marketing departments invest heavily in by way of search engine optimization (Google, 2020). Approximately half the domains included in

this thesis were for-profit U.S. entities; found using Google (which was used exclusively in this thesis). Because of this, regions other than New Zealand and Australia (“West Pacific”) were necessarily merged into a single category (“World/Other”) in order to for statistical comparisons to proceed. No doubt a larger study using other popular search engines would go a long way in overcoming this. Admittedly, the present author may have chosen a slightly different approach in overcoming this limitation through use of descriptive statistics. However this was considered suboptimal because such approach would preclude formal hypothesis testing (see, Section 1.7).

A second limiting factor concerned use of the DISCERN instrument. In this thesis comparisons (between locations and organisation type) were based on item #16. Charnock (1998, p.41) devised this as an “intuitive summary” of items #1–15; and indeed this is how the DISCERN instrument has been implemented (e.g., Alsoghier et al., 2018; Grewal and Alagaratnam, 2013b; Wiriyaakijja et al., 2016; Manchaiah et al., 2019). The problem with doing so in this thesis was that it constrained the dependent variable to that of a discrete, ordinal measure where other variables were continuous, e.g., PEMAT understandability, Plain Language, SAM+CAM. Such an approach rendered DISCERN scores inappropriate for omnibus analyses, otherwise adopted to better control for Type I statistical errors (incorrect rejection of a true null hypothesis H_0). Presumably to resolve this, some authors choose to report the sum total of items #1–16 or derive a percentage from this out of the maximum possible score of $16 \times 5 = 80$ (e.g., Seymour et al., 2015b; Alamoudi and Hong, 2015b; Doruk et al., 2018; Jo et al., 2018; Narwani et al., 2016). Here, a similar attempt was made, albeit in post-hoc fashion, by dividing the summary score by 5 and multiplying the result by 100. This at least allowed comparisons to be made with other studies using a genuine continuous scale; so long as it was accepted than comparisons would be inaccurate to $\pm 20\%$ ¹.

Third, In terms of the approach taken here with keyword searches, this

¹Minimum percentage step size necessarily equals 20% because a five point scale expressed as a percentage is $100/5=20$)

thesis relied exclusively on Google despite there being a number of alternative and popular search engines, such as DuckDuckGo, Bing, Yahoo, Ask, and emerging privacy conscious search engines StartPage, Qwant, and searX. This raises questions as to whether or not top ranking web materials on implantable hearing devices were representative of how people actually research this topic. That said, at the time of writing Google does account for three-quarters of desktop searches globally, higher still for mobile search traffic (99Firms.com, 2019).

Fourth, overall, the PEMAT tool proved useful in this thesis because it permitted audiovisual materials to be analysed along the same lines as print materials; minus numeric literacy assessment (c.f. Tables 2.1 and 2.2). The reason this is being mentioned as a limitation is not so much a problem with the PEMAT-AV tool so much as it was a problem in this thesis due to lack of regional granularity, mentioned above. Except here it proved more limiting; firstly, all domains were located in the Americas and so regional comparisons could not be made. Second, neither of the two domains hosting the videos (YouTube, Khan Academy) were HONcode certified. So no HONcode comparison could be made either. Third, eight of nine videos were hosted by a single for-profit entity (YouTube), while only one was classified as non-for-profit (Khan Academy) and so entity-type comparisons were also not possible. Again, a larger sample size yielded from Google, Bing, Yahoo may well address this.

Fifth, by design, no keywords beyond those earlier outlined in Section 2.2.1 were excluded prior to internet searches. “Auditory brainstem implants,” for instance, were not excluded, despite the fact this procedure is not widely known yet in New Zealand¹ and certainly is not yet currently funded in any systematic way by the Ministry of Health (as of 2019). Yet the study participants were all New Zealand based and therefore reflect information-gathering habits aligning more closely to the

¹Neurotologist Michel Neef has specific training in this procedure. To the best of this author’s knowledge, he is the only surgeon in New Zealand qualified to perform auditory brainstem implants. Readers are referred to following clickable links:

www.entassociates.co.nz

www.scoop.co.nz

sorts of treatment options offered in New Zealand. Because of this, it is difficult to draw robust conclusions about readability and suitability results from webpages located in the “Americas” compared to those of “West Pacific” countries in light of the recruitment process. Wider reach, e.g., through FaceBook paid advertising, is one way this could be overcome.

Lastly, concerning that of the SMOG index, (Mc Laughlin, 1969) asserts that 100-word sampling, as applied here for all readability measures, is so small as to be “uncharacteristic of the text being assessed,” (Mc Laughlin, 1969, p.641). In the context of today’s somewhat terse health-related writing, it is not clear whether 600 words may have changed our results appreciably compared to using 100 words. Most if not all of the readability studies cited in this thesis mention 100 word sampling. So, at least for comparability purposes there would still appear to be validity in this approach. This point is only raised so readers are aware of original validation of the SMOG tool.

4.8 Future recommendations

The following are recommended for any future readability and suitability research relating to implantable hearing devices:

1. Recruitment should leverage social media advertising for global reach. FaceBook would be ideal (similar to methods used by Serban (2018), a previous MAud student within the same group). Suggestions for ad targeting: (1) choose “interests” and “behaviors” based on or related as closely as possible to relevant keywords, such as “cochlear implant” and “BAHA” etc; (2) target specific geographic areas; (3) target parents.
2. Include low-volume search terms. For instance, entering the following term in Google Trends: “auditory brainstem implant - Topic” yielded a total of 1,800 worldwide searches within the past 12 months, ended 22 January 2020.

3. Use more than one search engine for compiling a list of webpages.
Yahoo and Bing are often cited in studies along with Google.
Privacy-focussed browsers such as DuckDuckGo, StartPage, Qwant, searX, are gaining popularity—perhaps it would be useful to include one of these (the author suggests StartPage, or DuckDuckGo) in addition to the more mainstream search engines above.
4. Concerning statistical analyses, in particular where continuous variables are used from other tools, express DISCERN results as a percentage.

Chapter 5

Conclusion

With respect to implantable hearing devices, the following conclusions are drawn:

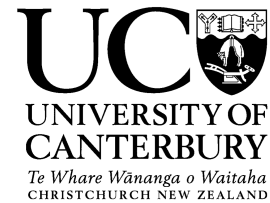
1. Only 5% of websites from the present study met readability guidelines of $\leq 8^{th}$ RGL. This is consistent with findings from related studies. Therefore, website content designers, by and large, appear to be writing at far too technical a level for most lay adults to comprehend. This presents a significant barrier for people naive to implantable hearing device technology and surgery who wish to undertake their own due diligence before engaging in shared decision making with their healthcare provider. Since independent research would be a precursor to shared decision making, ultimately this may impact decisions regarding treatment.
2. Organisation type, namely for-profit and not-for-profit, do not appear to matter in terms of how understandable and how suitable webpages might be for consumers of health information. From a parent's point of view (i.e., those with deaf children), or an adult with profound deafness, not needing to mentally filter out commercial or government websites can only serve to maximise resources for information gathering, which is empowering.
3. Similarly, which geographic location information seekers reside in (i.e., presumed audiences) doesn't appear to have any effect on how

understandable or suitable webpages are for people whose first language is English. Health information consumers residing in New Zealand or Australia for instance would find materials published within the Americas (e.g., NIH, Johns Hopkins) just as comprehensible and just as suitable as content published on the Ministry of Health (New Zealand), or Cochlea (Australia).

4. This study found no evidence to suggest that HONcode certified websites may be deemed any more readable or suitable than those without certification. Therefore, it is quite unlikely that health information seekers would even notice whether a website is HONcode certified or not; and even less likely that this would influence trust and browsing behaviour.

Appendix

Attached are two supporting documents. The first is to the Chain of Human Ethics Committee, regarding approval for this project (this author's name was subsequently appended by email communication—not shown). The second is a copy of the survey brief issued to potential survey participants.



HUMAN ETHICS COMMITTEE

Secretary, Rebecca Robinson
Telephone: +64 03 369 4588, Extn 94588
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2019/07/LR

8 April 2019

Ana Blagojevic, Aynsley Hickson, Carol Hewitt, Katie Murphy, and Sarah Folkerts
Psychology, Speech and Hearing
UNIVERSITY OF CANTERBURY

Dear Ana, Aynsley, Carol, Katie, and Sarah

Thank you for your request for an amendment to your research proposal "Quality of Hearing-Related Internet Information" as outlined in your email dated 2nd April 2019.

I am pleased to advise that this request has been considered and approved by the Human Ethics Committee.

Yours sincerely

A handwritten signature in black ink, appearing to be 'DS' followed by a stylized flourish.

Dr Dean Sutherland
Chair, Human Ethics Committee

Quality of Hearing-Related Internet Information

Matthew Goddard

Master of Audiology Student (2nd year)

matthew.goddard@pg.canterbury.ac.nz

Tel. 022 0964129

SUPERVISOR: REBECCA KELLY-CAMPBELL
ASSOCIATE SUPERVISOR: MEGAN MCAULIFFE

Department of Communication Disorders |
Tumuaki Tuarua, Te Tari Mātai Hauora Reo

University of Canterbury |
Te Whare Wānanga o Waitaha |
Private Bag 4800 | Christchurch 8140 | New Zealand

May 27, 2019



WHAT THIS PROJECT IS ALL ABOUT

What does the project seek to do?

This project is all about understanding how useful hearing-related web-information is for people. We do this by looking at reading grade level, use of visual aids, logical structure, use of jargon, and so on. Factors such as these influence how readable and how suitable webpages are for us whenever we type something in to Google. For instance, we might search "treatment for tinnitus" and click one or more webpages Google thinks is relevant for us. The trouble with that is, Google (and other search engines) have no way of knowing how well people understand what they read on those pages. What we plan on doing is measuring how readable and suitable some of these webpages are for people specifically interested hearing disorders and hearing technology (see 1–7 in list below).

We need your help

In order for us to know which pages to look at, we need your help. We need you to think of words you would type into Google as though you were genuinely interested in "implantable hearing devices."

How did this project arise?

People seek health information from a variety of sources. One source that is increasing in use is the internet. In fact, around 80% of internet users in the US search the internet for health information (Fox, 2006). Because people using the internet to find health information often do so without the guidance of a healthcare professional, it is important that internet health information is easy to read and understand and that it provides people with information to help them make informed decisions about their health.

In 2015, Laplante-Lévesque and Thorén (2015) conducted a systematic review of the readability of online hearing-related information in English and found that information was of variable readability, but generally the information exceeded the reading level recommended by the American Medical Association. To date, little published research exists examining the readability and quality of internet information about specific hearing disorders. This project aims to fill that research gap. Specifically, we are investigating readability and quality of information on the following:

1. Balance/Dizziness
2. Auditory Processing Disorder
3. Sudden Hearing Loss
4. Auditory Neuropathy
5. Middle Ear infections
6. Single-Sided Deafness
7. Implantable hearing devices

THINGS YOU NEED TO KNOW

As a participant in this *Quality of Hearing-Related Internet Information* study, conducted within the University of Canterbury (UC), we have a responsibility to ensure that any details you share with us ("UC") remain secure. To this end:

WE WILL NOT:

- × withhold information about the nature of your involvement in this study (you are welcome to email or telephone if you have any questions)
- × attempt to invade your privacy, electronically or otherwise
- × gather any personal information that is sensitive (we won't even ask for your name)
- × share your email address, if you choose to provide one (see below)

WHAT WE DO COLLECT

- ✓ basic demographic information (gender, English language fluency, level of education, ethnicity, age)
- ✓ any search terms you are able to supply us (we ask for at least three)
- ✓ your email address - this is completely optional and only required in order for us to send you a summary of research outcomes, once these become available. Your email address will be stored electronically in accordance with UC Data Management and Human Ethics policies, and will be deleted from our file servers after a period of five years from the date of your participation in this study.

INSTRUCTIONS FOR PARTICIPATION

If you're happy to participate in this study, the first thing you need to do is give your consent—you can do this by checking a tick-box on the same page you downloaded this form (first page). Once you've done this you can begin with the six demographic questions we ask, followed by the question on search terms.

Please note that we do require three or more search terms. We also encourage you to brainstorm search terms without typing them into Google and doing an actual search. This is because your prior browsing history could in principle influence the results you see. So please just "imagine" what you would type, then tell us what those search terms are for Q.7.

Your email address is the last thing we ask for. It is entirely up to you whether you want to supply this. It's just for us to send you a copy of the results of the study. Beyond this no further correspondence will be sent from us to you.

If you have any further questions about the study itself or would like to ask us about any concerns you might have before you participate, you are absolutely welcome to get in touch either by email text or phone (see first page for contact details).

REFERENCES

- Fox, S. (2006). Online health search 2006. Technical report, Pew Research Center: Internet and Technology.
- Laplante-Lévesque, A. and Thorén, E. S. (2015). Readability of internet information on hearing: Systematic literature review. *American Journal of Audiology*, 24(3):284–288.

References

- 99Firms.com (2019). Search engine statistics.
- Alamoudi, U. and Hong, P. (2015a). Readability and quality assessment of websites related to microtia and aural atresia. *International Journal of Pediatric Otorhinolaryngology*, 79(2):151–156.
- Alamoudi, U. and Hong, P. (2015b). Readability and quality assessment of websites related to microtia and aural atresia. *International Journal of Pediatric Otorhinolaryngology*, 79(2):151–156.
- Alsoghier, A., Ni Riordain, R., Fedele, S., and Porter, S. (2018). Web-based information on oral dysplasia and precancer of the mouth –quality and readability. *Oral Oncology*, 82:69–74.
- Amante, D. J., Hogan, T. P., Pagoto, S. L., English, T. M., and Lapane, K. L. (2015). Access to care and use of the internet to search for health information: results from the us national health interview survey. *J Med Internet Res*, 17(4):e106.
- Andreassen, H. K., Bujnowska-Fedak, M. M., Chronaki, C. E., Dumitru, R. C., Pudule, I., Santana, S., Voss, H., and Wynn, R. (2007). European citizens’ use of e-health services: A study of seven countries. *BMC Public Health*, 7(1):53.
- Begeny, J. C. and Greene, D. J. (2014). Can readability formulas be used to successfully gauge difficulty of ready materials? *Psychology in the Schools*, 51(2):198–215.
- Bellon-Harn, M., Manchaiah, V., and Morris, L. (2020). A cross-sectional

- descriptive analysis of portrayal of autism spectrum disorders in youtube videos: A short report. *Autism*, 24(1):263–268.
- Bento, R. F., Kiesewetter, A., Ikari, L. S., and Brito, R. (2012). Bone-anchored hearing aid (baha): indications, functional results, and comparison with reconstructive surgery of the ear. *International archives of otorhinolaryngology*, 16(3):400–405.
- Berkman, N. D., Davis, T. C., and McCormack, L. (2010). Health literacy: What is it? *Journal of Health Communication*, 15(sup2):9–19. PMID: 20845189.
- Berners-Lee, T., Cailliau, R., Groff, J., and Pollermann, B. (2010). World-wide web: the information universe. *Internet Research*, 20(4):461–471.
- Boston, M., Ruwe, E., Hilbert, L., Choo, D., and Greinwald, J. (2004). Cochlear implant information on the internet. *International Congress Series*, 1273:413–416.
- Boyer, C., Baujard, V., and Geissbuhler, A. (2011). Evolution of health web certification through the honcode experience. *Studies in health technology and informatics*, 169:53.
- Boyer, C. and Dolamic, L. (2015). Automated detection of honcode website conformity compared to manual detection: an evaluation. *Journal of Medical Internet Research*, 17(6):e135.
- Carroll, J. B. (1987). The national assessments in reading: Are we misreading the findings? *The Phi Delta Kappan*, 68(6):424–430.
- Castro-Sánchez, E., Spanoudakis, E., and Holmes, A. H. (2015). Readability of ebola information on websites of public health agencies, united states, united kingdom, canada, australia, and europe. *Emerging infectious diseases*, 21(7):1217–1219.
- Charnock, D. (1998). *The DISCERN Handbook: Quality criteria for consumer health information on treatment choices*. Radcliffe Medical Press.

- Charnock, D., Shepperd, S., Needham, G., and Gann, R. (1999). Discern: an instrument for judging the quality of written consumer health information on treatment choices. *Journal of epidemiology and community health*, 53(2):105–111.
- Chu, J. T., Wang, M. P., Shen, C., Viswanath, K., Lam, T. H., and Chan, S. S. C. (2017). How, when and why people seek health information online: Qualitative study in hong kong. *Interactive journal of medical research*, 6(2):e24–e24.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297.
- de Castro Correa, C., Ferrari, D., and Berretin-Felix, G. (2013). Quality, range, and legibility in web sites related to orofacial functions. *International archives of otorhinolaryngology*, 7(4):358–362.
- Dewalt, D., Berkman, N., Sheridan, S., Lohr, K., and Pignone, M. (2004). Literacy and health outcomes: a systematic review of the literature. *Journal of General Internal Medicine*, 9(12):1228–1239.
- Doak, C., Doak, L., and Root, J. (1996). *Teaching Patients with Low Literacy Skills*, chapter 4, pages 49–60. J. B. Lippincott Company Philadelphia, second edition.
- Dobrev, I., Sim, J. H., Stenfelt, S., Ihrle, S., Gerig, R., Pfiffner, F., Eiber, A., Huber, A. M., and Rösli, C. (2017). Sound wave propagation on the human skull surface with bone conduction stimulation. *Hearing Research*, 355:1–13.
- Doruk, C., Enver, N., Çaytemel, B., Azezli, E., and Başaran, B. (2018). Readability, understandability, and quality of online education materials for vocal fold nodules. *Journal of Voice*.
- Dreyzehner, J. and Goldberg, K. A. (2019). Depression in deaf and hard of hearing youth. *Child and Adolescent Psychiatric Clinics of North America*, 28(3):411–419.

- Feiner, D. E. and Rayan, G. M. (2016). Preoperative surgical discussion and information retention by patients. *Journal of Hand Surgery (Asian-Pacific Volume)*, 21(3):326–332.
- Fellinger, J., Holzinger, D., Sattel, H., Laucht, M., and Goldberg, D. (2009). Correlates of mental health disorders among children with hearing impairments. *Developmental Medicine and Child Neurology*, 51(8):635–641.
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, 32(3):221–233.
- Fox S, D. M. (2012). Mobile health 2012. *Pew Research Center: Internet and Technology*.
- Fry, E. (1979). *Elementary reading instruction*. McGraw-Hill Companies.
- Glenton, C., Paulsen, E. J., and Oxman, A. D. (2005). Portals to wonderland: Health portals lead to confusing information about the effects of health care. *BMC Medical Informatics and Decision Making*, 5(1):7.
- Glickman, N. S. (01.01.2013). Lessons learned from 23 years of a deaf psychiatric inpatient unit.
- Google (2020). Search engine optimization (seo) starter guide.
- Greer, C. (2019). *Readability and quality of online information regarding tinnitus in the Spanish language*. PhD thesis, University of Canterbury, School of Psychology — Speech and Hearing.
- Greene, M., Cleary, Y., and Marcus-Quinn, A. (2017). Use of plain-language guidelines to promote health literacy. *IEEE Transactions on Professional Communication*, 60(4):384–400.
- Grewal, P. and Alagaratnam, S. (2013a). The quality and readability of colorectal cancer information on the internet. *International Journal of Surgery*, 11(5):410–413.

- Grewal, P. and Alagaratnam, S. (2013b). The quality and readability of colorectal cancer information on the internet. *International Journal of Surgery*, 11(5):410–413.
- Gunn, R. G. (2010). Adult cochlear implants in new zealand—a chronic funding issue. *The New Zealand Medical Journal*, 123(1321):9–12.
- Gunning, R. (1952). *The Technique of Clear Writing*. McGraw-Hill.
- Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, 6(2):3–13.
- Harris, V. C., Links, A. R., Hong, P., Walsh, J., Schoo, D. P., Tunkel, D. E., Stewart, C. M., and Boss, E. F. (2018). Consulting dr. google: Quality of online resources about tympanostomy tube placement. *The Laryngoscope*, 128(2):496–501.
- Helitzer, D., Hollis, C., Cotner, J., and Oestreicher, N. (2009). Health literacy demands of written health information materials: An assessment of cervical cancer prevention materials. *Cancer Control*, 16(1):70–78.
- HONcode (2020). Health on the net foundation.
- Ingram, D. (2015). Percentage of u.s. adults who looked up health information on the internet in the past year, by type of locality —. *National Health Interview Survey*, 64(49):1367.
- InternetLiveStats (2018). Total number of websites. *The official World Wide Web Anniversary Site: World Wide Web Foundation*, June, www.internetlivestats.com.
- IWS (2019). World internet users and 2019 poulation stats. *Miniwatts Marketing Group*, Retrieved from: www.internetworldstats.com.
- Jadhav, A., Andrews, D., Fiksdal, A., Kumbamu, A., McCormick, J. B., Misitano, A., Nelsen, L., Ryu, E., Sheth, A., Wu, S., and Pathak, J. (2014). Comparative analysis of online health queries originating from personal computers and smart devices on a consumer health information portal. *J Med Internet Res*, 16(7):e160.

Jo, J. H., Kim, E. J., Kim, J. R., Kim, M. J., Chung, J. W., and Park, J. W. (2018). Quality and readability of internet-based information on halitosis. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, 125(3):215–222.

Joseph, J., Svider, P. F., Shaigany, K., Eloy, J. A., McDonald, P. G., Folbe, A. J., and Hong, R. S. (2016). Hearing aid patient education materials: Is there room for improvement? *J Am Acad Audiol*, 27(4):354–359.

Joseph L. Fleiss, B. L. and Paik, M. C. (2003). *Statistical Methods for Rates and Proportions*, chapter The measurement of interrater agreement, pages 598–626. Number 18 in Wiley Series in Probability and Statistics. John Wiley and Sons Inc., 3rd edition.

Kaestle, C., Campbell, A., Finn, J., Johnson, S., and Mikulecky, L. (2001). Adult literacy and education in america: four studies based on the national adult literacy survey. *U.S. Department of Education Office of Educational Research and Improvement*, NCES 2001–534.

Kang, S. J. and Lee, M. S. (2019). Assessing of the audiovisual patient educational materials on diabetes care with pemat. *Public Health Nursing*, 36(3):379–387.

Kelley, N. H. (1937). Historical aspects of bone conduction. *The Laryngoscope*, 47(2):102–109.

Kim, H., Paige Powell, M., Bhuyan, S. S., and Bhuyan, S. S. (2017). Seeking medical information using mobile apps and the internet: Are family caregivers different from the general public? *J Med Syst*, 41(3):38.

Kincaid, P., Jr., R. F., Rogers, R., and Chissom, B. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. *Institute for Simulation and Training*, 56.

- Kirby, P., Reynolds, K., Walker, J., Furer, P., and Pryor, T. (2018). Evaluating the quality of perinatal anxiety information available online. *Archives of Women's Mental Health*, 21(6):813–820.
- Kirsch, I. S., Jungebut, A., Jenkins, L., and Kolstad, A. (1993). *Adult literacy in America: a first look at the results of the National Adult Literacy Survey*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Kortlever, J., Derkzen, L., Kleiss, I., Reichel, L., and Vagner, G. (2019). Quality and readability of online information on carpal tunnel syndrome. *Journal of Hand Surgery (European Volume)*, 44(9):979–980.
- Kutner, M., Greenberg, E., Jin, Y., and Paulsen, C. (2003). The health literacy of america's adults: Results from the 2003 national assessment of adult literacy (nces 2006–483). Technical report, U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Kvam, M. H., Loeb, M., and Tambs, K. (2007). Mental health in deaf adults: Symptoms of anxiety and depression among hearing and deaf individuals. *Journal of Deaf Studies and Deaf Education*, 12(1):1–7.
- Laplante-Lévesque, A. and Thorén, E. S. (2015). Readability of internet information on hearing: Systematic literature review. *American Journal of Audiology*, 24(3):284–288.
- Lassaletta, L., Sánchez-Cuadrado, I., Espinosa, J. M., Batuecas, Á., Cenjor, C., Lavilla, M. J., Cavallé, L., Huarte, A., Nuñez, F., Manrique, M., Ramos, Á., de Paula, C., and Gil-Carcedo, E. (2019). Active middle ear implants. *Acta Otorrinolaringologica (English Edition)*, 70(2):112–118.
- Laurent, M. R. and Vickers, T. J. (2009). Seeking health information online: Does wikipedia matter? *Journal of the American Medical Informatics Association*, 16(4):471–479.
- Lee, Y.-M., Yu, H. Y., You, M.-A., and Son, Y.-J. (2017). Impact of health literacy on medication adherence in older people with chronic diseases. *Collegian*, 24(1):11–18.

- Macherey, O. and Carlyon, R. P. (2014). Cochlear implants. *Current Biology*, 24(18):R878–R884.
- Manchaiah, V., Dockens, A. L., Flagge, A., and Bellon-Harn, M. (2019). Quality and readability of english-language internet information for tinnitus. *Journal of the American Academy of Audiology*, 30(1).
- Mayer, G. G. and Villaire, M. (2007). *Health literacy in primary care: a clinician's guide*. Springer Publishing Company, New York.
- Mc Laughlin, G. H. (1969). Smog grading-a new readability formula. *Journal of Reading*, 12(8):639–646.
- Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education*, 30(4):159–167. PMID: 17108243.
- MoE (2008). Literacy and numeracy in new zealand: findings from the adult literacy and life skills survey. *Ministry of Education, New Zealand Government*,
<https://www.educationcounts.govt.nz/publications/series/Adult-Literacy>.
- MoH (2010). Korero marama: Health literacy and maori results from the 2006 adult literacy and life skills survey. *Wellington: Ministry of Health*.
- MoH (2019). Health literacy. *Making Services Better for Users*, Retrieved from: www.health.govt.nz/our-work/making-services-better-users/health-literacy.
- Moore, B. and Alcántara, J. (2001). The use of psychophysical tuning curves to explore dead regions in the cochlea. *Ear and Hearing*, 22(4):268–278.
- Moore, B. and Glasberg, B. (1997). A model of loudness perception applied to cochlear hearing loss. *Auditory Neuroscience*, 3:289–311.
- Nakatomi, H., Miyawaki, S., Kin, T., and Saito, N. (2016). Hearing restoration with auditory brainstem implant. *Neurologia medico-chirurgica*, 56(10):597–604.

- Narwani, V., Nalamada, K., Lee, M., Kothari, P., and Lakhani, R. (2016). Readability and quality assessment of internet-based patient education materials related to laryngeal cancer. *Head & Neck*, 38(4):601–605.
- NCES (2019). *Reading performance*. Number Chapter 1/Preprimary, Elementary, and Secondary Education in Section: Assessments. National Center for Education Statistics.
- NIH (2019). Cochlear implants. *National Institute on Deafness and Other Communication Disorders (NIDCD)*, Retrieved from: <https://www.nidcd.nih.gov/health/cochlear-implants>.
- Nutbeam, D. (2000). Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century. *Health promotion international*, 15(3):259–267.
- Nutbeam, D. (2008). The evolving concept of health literacy. *Social Science and Medicine*, 67(12):2072–2078.
- Online-Utility.org (2009). Tests document readability: readability calculator www.online-utility.org.
- Parikh, N., Parker, R., Nurss, J., Baker, D., and Williams, M. (1996). Shame and health literacy: the unspoken connection. *Patient Education and Counseling*, 27(1):33–39.
- Peddie, K. A. and Kelly-Campbell, R. J. (01.08.2017). How people with hearing impairment in new zealand use the internet to obtain information about their hearing health. *Computers in human behavior*, 73:141–151.
- Porter, A. and Edirippulige (2007). Parents of deaf children seeking hearing loss-related information on the internet: the australian experience. *Journal of Deaf Studies and Deaf Education*, 12(4):518–529.
- Redish, J. (1981). Understanding the limitations of readability formulas. *IEEE Transactions on Professional Communication*, PC-24(1):46–48.
- Redish, J. (2000). Readability formulas have even more limitations than klare discusses. *ACM J. Comput. Doc.*, 24(3):132–137.

- Rello, L. and Baeza-Yates, R. (2017). How to present more readable text for people with dyslexia. *Universal Access in the Information Society*, 16(1):29–49.
- Ritchie, L., Tornari, C., Patel, P., and Lakhani, R. (2016). Glue ear: how good is the information on the world wide web? *The Journal of Laryngology and Otology*, 130(2):157–161.
- Robins, S., Barr, H. J., Idelson, R., and Lambert, S. (21.10.2016). Online health information regarding male infertility: An evaluation of readability, suitability, and quality. *Interactive journal of medical research*, 5(4).
- Ruble, K. J., Paré-Blagoev, E. J., Cooper, S. L., and Jacobson, L. A. (2019). Assessment of online resources for returning to school during and after treatment of childhood cancer. *Journal of Cancer Education*.
- Salama, A., Panoch, J., Bandali, E., Carroll, A., Wiehe, S., Downs, S., Cain, M. P., Frankel, R., and Chan, K. H. (2019). Consulting “dr. youtube”: an objective evaluation of hypospadias videos on a popular video-sharing website. *Journal of Pediatric Urology*.
- Santana, S., Lausen, B., Bujnowska-Fedak, M., Chronaki, C. E., Prokosch, H.-U., and Wynn, R. (2011). Informed citizen and empowered citizen in health: results from an european survey. *BMC Family Practice*, 12(1):20.
- Saroul, N., Gilain, L., Montalban, A., Giraudet, F., Avan, P., and Mom, T. (2011). Patient satisfaction and functional results with the bone-anchored hearing aid (baha). *European Annals of Otorhinolaryngology, Head and Neck Diseases*, 128(3):107–113.
- Sax, L., Razak, A., Shetty, K., Cohen, M., and Levi, J. (2019). Readability of online patient education materials for parents after a failed newborn hearing screen. *International Journal of Pediatric Otorhinolaryngology*, 125:168–174.

Selden, C., Zorn, M., Ratzan, S., and Parker, R. (2000). National library of medicine current bibliographies in medicine: Health literacy.

Bethesda (MD): National Library of Medicine, page vi.

Serban, M. (2018). *A readability analysis of French language online information on hearing related websites*. PhD thesis, University of Canterbury, Department of Communication Disorders.

Seymour, N., Lakhani, R., Hartley, B., and Cochrane, L. (2015a).

Cochlear implantation: An assessment of quality and readability of web-based information aimed at patients. *Cochlear implants international*, 16(6):321–325.

Seymour, N., Lakhani, R., Hartley, B., and Cochrane, L. (2015b).

Cochlear implantation: An assessment of quality and readability of web-based information aimed at patients. *Cochlear implants international*, 16(6):321–325.

Shoemaker, S., Wolf, M., and Brach, C. (2013a). Patient education materials assessment tool for audiovisual materials (pemat-a/v).

Prepared by Abt Associates, under Contract No. HHSA290200900012I, TO 4). Rockville, MD: Agency for Healthcare Research and Quality; AHRQ Publication No. 14-0002-EF.

Shoemaker, S., Wolf, M., and Brach, C. (2013b). Patient education

materials assessment tool for printable materials (pemat-p). *Prepared by Abt Associates under Contract No. HHSA290200900012I, TO 4). Rockville, MD: Agency for Healthcare Research and Quality; AHRQ Publication No. 14-0002-EF.*

Shoemaker, S., Wolf, M., and Brach, C. (2013c). The patient education

materials assessment tool (pemat) and user's guide: an instrument to assess the understandability and actionability of print and audiovisual education materials. *Agency for Healthcare Research and Quality U.S. Department of Health and Human Services*, 14-0002-EF.

Shoemaker, S. J., Wolf, M. S., and Brach, C. (2014). Development of the patient education materials assessment tool (pemat): A new measure of

- understandability and actionability for print and audiovisual patient information. *Patient Education and Counseling*, 96(3):395–403.
- Shrout, P. E., Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological bulletin*, 86(2):420–428.
- Sillence, E., Briggs, P., Harris, P., and Fishwick, L. (2007). Going online for health advice: Changes in usage and trust practices over the last five years. *Interacting with computers*, 19(3):397–406.
- Sladen, D. P. and Zappler, A. (2015). Older and younger adult cochlear implant users: Speech recognition in quiet and noise, quality of life, and music perception. *American Journal of Audiology*, 24:31–39.
- Snik, A., Arjan, B., Mylanus, E., and Cremers, C. (2004). Candidacy for the bone-anchored hearing aid. *Audiology and Neurotology*, 9(4):190–196.
- Thorndike, E. L. (1921). *The Teacher's Work Book*. New York Teachers College, Columbia University.
- Tisch, M. (2017). Implantable hearing devices. *GMS current topics in otorhinolaryngology, head and neck surgery*, 16:Doc06; Doc06–Doc06.
- Toms, E. and C, L. (2007). How consumers search for health information. *Health informatics journal*, 13(3):223–235.
- Van Gent, T., Goedhart, A. W., Hindley, P. A., and Treffers, P. D. A. (2007). Prevalence and correlates of psychopathology in a sample of deaf adolescents. *Journal of Child Psychology and Psychiatry*, 48(9):950–958.
- Wangberg, S. (01.12.2009). Use of the internet for health purposes: trends in norway 2000-2010. *Scandinavian journal of caring sciences*, 23(4):691–696.
- Weiss, B. D. (2003). *Health literacy: A manual for clinicians*, volume Part of an educational program about health literacy. American Medical Association.

WHO (2019). Deafness and hearing loss. *Fact sheets*, Retrieved from:
www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss.

Wiriyakijja, P., Fedele, S., Porter, S., and Ni Riordain, R. (2016).
Web-based information on the treatment of oral leukoplakia –quality
and readability. *Journal of Oral Pathology & Medicine*, 45(8):617–620.

Zamanian, M. and Heydari, P. (2012). Readability of texts: state of the
art. *Theory and Practice in Language Studies*, 2(1):43–53.